Euclid's Algorithm for the Greatest Common Divisor

Desh Ranjan
Department of Computer Science
New Mexico State University

1 Numbers, Division and Euclid

Mathematician (to shepherd): Hey, you have exactly 137 sheep in your flock, right?

Shepherd: Yes, indeed! how did you count them so fast?

Mathematician: It's easy, really. I just counted their legs and divided that by four.

It should not surprise you that people have been using numbers (and operations on them like division) for a very long time for very practical purposes like dividing up the money left by parents equally among children or distributing ears of corn equally to groups of people and more generally to conduct all sorts of business dealings. It may be a bit of surprise to some of you that things like calculating divisors of numbers actually form the very core of methods that ensure security of all computer systems and internet communications. The RSA cryptosystem that is used extensively for secure communications is based on the assumptions of difficulty of calculating divisors of large numbers. So, people have been thinking of divisors of numbers for a very, very long time and calculating these actually has significant import in the modern day world. A very related and even more basic notion is the notion of multiples of quantities. One very natural way to compare quantities is to "measure" how many times do we need to aggregate the smaller quantity to obtain the larger quantity. For example, we may be able to compare two unknown lengths by observing that the larger length can be obtained by "aggregating" the smaller length three times. This provides one a sense of how the two lengths compare without actually knowing the two lengths. Unfortunately, the larger quantity may not always be obtainable from the smaller quantity by aggregating it an integral number of times. In this scenario, one way to think would be to imagine each of the two quantities to

be made up of smaller (identical) parts such that both the quantities can be obtained by aggregating these smaller parts an integral number of times. Obviously, for the larger quantity we will have to use a greater number of parts than the smaller quantity. For example, when comparing two weights, one might observe that larger one can be obtained by aggregating some weight 7 times whereas the smaller weight can be obtained by aggregating the same weight 5 times. This provides us a sense of how big the first weight is to the second one. Of course, in the above scenario, one can also observe that if we chose even smaller parts to "split" the weights (say a quarter of the first one) the first weight would be we obtained by aggregating this even smaller weight 28 times and the smaller of the two original weights would be obtained by aggregating this smaller part 20 times which also provides us a sense of relative magnitudes of the two weights. However, having smaller numbers like 7 and 5 to describe relative magnitudes seems intuitively and practically more appealing as opposed to 28 and 20. This leads us to think about what would be the greatest magnitude such that two given magnitudes will both be multiples of that common magnitude. It turns out that this is a question that was given some thought by Greek mathematicians more than 2000 years ago. An important figure in (Greek) mathematics by the name of Euclid was one of the first to compile a collection of mathematical works called *Elements* that has a chapter, interestingly called a "Book", about numbers. During the course of this project you will read a translation of part of this chapter to discover the Euclid's method(algorithm) to compute the greatest common divisor of two numbers. It is not clear if Euclid was the first person to discover this algorithm or just the first to record it in writing.

1.1 Who was Euclid?

One of the most influential mathematicians of ancient Greece, Euclid, lived around 300 B.C.E. Very little is known about his life. It is genarally believed that he was educated at Plato's academy in Athens, Greece. According to Proclus (410-485 C.E.) Euclid came after the first pupils of Plato and lived during the reign of Ptolemy I (306-283 B.C.E.). It is said that Euclid established a mathematical school in Alexandria. By all accounts, it seems that he was a kind, fair, patient man. Euclid is best known for his mathematical compilation *Elements* in which among other things he laid down the foundations of geometry and number theory. The geometry that we learn in school today traces its roots to this book and Euclid is sometimes called the father

of geometry. Euclid did not study mathematics for its potential practical applications or financial gains. He studied mathematics for a sense of order, structure and the ideal form of reason. To him geometrical objects and numbers were abstract entities and he was interested in studying and discovering their properties. In that sense, he studied mathematics for its own sake. One story that reveals his disdain for learning for the purpose of material gains concerns a pupil that had just finished his first geometry lesson. The pupil asked what he would gain from learning geometry. As the story goes, Euclid asked his subordinate to give the pupil a coin so that he would be gaining from his studies. Another story that reveals something about his character concerns King Ptolemy. Ptomlemy asked the mathematician if there was an easier way to learn geometry. Euclid replied, "there is no royal road to geometry", and sent the king to study.

Euclid wrote several books like (*Data, On Divisions of Figures, Phaenomena, Optics*, and the lost books *Conics* and *Porisms*), but *Elements* remains his best known compilation. The first "book" in this compilation is perhaps the most well-known. It lays down the foundations of what we today call "Euclidean" geometry (which was the only geometry people studied until the Renaissance). This book has the definitions of basic geometric objects like points and lines along with the basic postulates or axioms. These axioms are then used by Euclid to establish many other truths (*Theorems*) of geometry. *Elements* is considered one of the greatest works of mathematics. It was translated into Latin and Arabic and influenced mathematics throughout Europe and the middle east. It was probably the standard "textbook" for geometry for more than 1500 years in western Europe and continues to influence the way geometry is taught to this day.

Book 7 of *Elements* provides foundations for number theory. Euclid's Algorithm for calculating the greatest common divisor of two numbers was presented in this book. As one will notice later, Euclid uses lines to represent numbers and often relies on visual figures to aid the explanation of his method of computing the GCD. As such, he seems to be relating numbers to geometry which is quite different from present day treatment of number theory.

Today, erroneously, many different methods are called Euclid's algorithm to compute the greatest common divisor(GCD) of two numbers. By reading the original writings of Euclid you will discover the real Euclid's algorithm and appreciate its subtlety and beauty. In any case, "Euclid's Algorithm" is one of the most cited and well-known examples of an (early) algorithm. To quote Knuth [1] :

By 1950, the word algorithm was mostly associated with "Euclid's Algorithm".

2 Prelude

We say that a number [1] x divides another number y if y is a multiple of x. For example, $1, 2$, and 3 all divide 6 but 5 does not divide 6. The only divisors of 17 are 1 and 17. The notation $x|y$ is a shorthand for "x divides y". We denote by $divisors(x)$ the set of all the numbers y such that $y|x$. So, for example, $divisors(6) = \{1, 2, 3, 6\}$ and $divisors(60) = \{1, 2, 3, 4, 5, 6, 10, 12, 15, 20, 30, 60\}$.

A number z is called a common divisor of two numbers x and y if $z|x$ and $z|y$. We denote by $cd(x, y)$ the set of all common divisors of x and y. For example, $cd(6, 8) = \{1, 2\}$ and $cd(40, 180) = \{1, 2, 4, 5, 10, 20\}$.

TASKS:

2.1 What is the set of divisors of the number 315?

2.2 Calculate the set $cd(420, 108)$.

2.3 Calculate the sets $cd(420, 216)$, and $cd(108, 96)$.

2.4 Prove that $cd(x, y) = divisors(x) \cap divisors(y)$.

While it is relatively easy to calculate the divisors of a number and common divisors of two numbers when the numbers are small, the task becomes harder as the numbers becomes larger.

TASKS:

2.5 Calculate $divisors(3456)$.

2.6 Calculate $cd(3456, 4563)$.

[1] The word number in this section means a positive integer. That is what it meant to Euclid.

2.7 Think about how you calculated $divisors(3456)$. Explain in your own words the method you used for this calculation.

2.8 Provide a step by step method to calculate the divisors of any given number x. Why does your method work? How many steps might your method require to calculate $divisors(x)$?

2.9 Write a JAVA/C program that uses your step by step method and computes $divisors(x)$ for any given number x.

2.10 Explain, in words, your method to calculate $cd(3456, 4563)$.

2.11 Provide a step by step method to calculate $cd(x, y)$ for any two numbers x and y. Why does this method work? How many steps might your method take to calculate $cd(x, y)$?

2.12 Write a JAVA/C Program that uses your step by step method to compute $cd(x, y)$ for any two given numbers x and y.

As you might have noticed the number 1 divides every number. Since there is no number smaller than 1, 1 is the **smallest** common divisor for any two numbers x and y. What about the **greatest** common divisor ? The greatest common divisor of two numbers x and y, denoted by $gcd(x, y)$, is the largest number z such that $z|x$ and $z|y$. Finding the greatest common divisor is not quite as easy as finding the smallest common divisor. But, by this time, you may have figured out a way to do so.

TASKS:

2.13 Prove that any two numbers x and y have a greatest common divisor.

2.14 Give a step by step method (algorithm) that given any two numbers x and y computes $gcd(x, y)$.

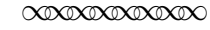
2.15 Use your method to calculate $gcd(1631, 1008)$. How many steps did your method need to calculate this gcd? In general, how many steps might your method take for computing $gcd(x, y)$.

2.16 Prove that your method always works correctly.

2.17 Write a JAVA/C program that uses your method to compute $gcd(x, y)$.

3 Euclid's Algorithm

Here we present the translations of (relevant) Definitions, Proposition 1 and Proposition 2 from Book VII of Euclid's *Elements* as translated by Sir Thomas L. Heath. ([2]). Euclid's method of computing the GCD is based on these propositions.



BOOK VII of *Elements* by Euclid

DEFINITIONS.

1. An **unit** is that by virtue of which each of the things that exist is called one.
2. A **number** is a multitude composed of units.
3. A number is **a part** of a number, the less of the greater, when it measures the greater.
4. but **parts** when it does not measure it. [2]
5. The greater number is a **multiple** of the less when it is measured by the less.
6. An **even number** is that which is divisible into two equal parts.
7. An **odd number** is that which is not divisible into two equal parts, or that differs by an unit from an even number.
8. An **even-times even number** is that which is measured by an even number according to an even number.
9. An **even-times odd number** is that which is measured by an even number according to an odd number.
10. An **odd-times odd number** is that which is measured by an odd number according to an odd number.

[2]While this definition is not relevant here, what is meant by this definition is quite subtle and subject of scholarly mathematical work

11. A **prime number** is that which is measured by an unit alone. [3]

12. Numbers **prime to one another** are those which are measured by an unit alone as a common measure.

13. A **composite number** is that which is measured by some number.

14. Numbers **composite to one another** are those which are measured by some number as a common measure.

TASKS:

3.1 Read these definitions carefully. You will notice various subtleties and will likely need to reread the definitions when answering further questions.

3.2 Discuss how Euclid's "unit" relates to the number 1. Does Euclid think that 1 is a number?

3.3 What does Euclid mean when he says that a number "measures" another number? Express Euclid's notion of "measures" in modern mathematical notation.

3.4 Does the number 4 measure number 72? Does 5 measure 72?

3.5 Euclid never defines what is a "common measure" but uses that in definition 12 and 14. What is your interpretation of Euclid's "common measure"?

3.6 Find a number (other than the unit) that is a common measure of numbers 102 and 187.

We now present Proposition 1 from Euclid's book VII. The proposition concerns numbers that are prime to one another.

[3] Reading further work of Euclid, e.g. Proposition 2, it is clear that Euclid meant that a prime number is that which is measured only by the unit and the number itself

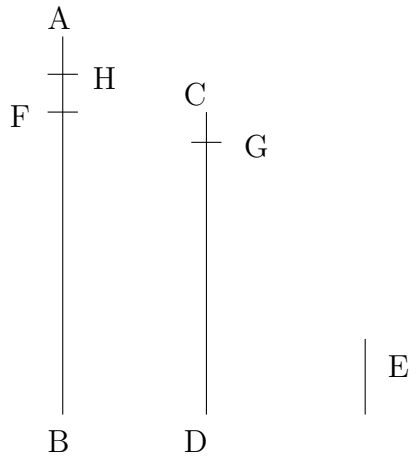
PROPOSITION 1.

Two unequal numbers being set out, and the less being continually subtracted in turn from the greater, if the number which is left never measures the one before it until an unit is left, the original numbers will be prime to one another.

For, the less of two unequal numbers AB, CD being continually subtracted from the greater, let the number which is left never measure the one before it until an unit is left;

I say that AB, CD are prime to one another, that is, that an unit alone measures AB, CD.

For, if AB, CD are not prime to one another, some number will measure them.



Let a number measure them, and let it be E; let CD, measuring BF, leave FA less than itself,

let, AF measuring DG, leave GC less than itself,

and let GC, measuring FH, leave an unit HA.

Since, then E measures CD, and CD measure BF, therefore E also measures BF.

But it also measures the whole BA;

therefore it will also measure the remainder AF.

But AF measures DG;

therefore E also measures DG.

But it also measures the whole DC;

therefore it will also measure the remainder CG.

But CG measures FH;

therefore E also measures FH.

But it also measures the whole FA;

therefore it will also measure the remainder, the unit AH, though it is a number: which is impossible.

Therefore no number will measure the numbers AB, CD; therefore AB, CD are prime to one another. [VII. Def 12]

Q. E. D.

TASKS:

3.7 Read the above proposition and its proof carefully to understand the process of repeated subtraction that Euclid is describing. The picture that "illustrates" the process is somewhat misleading in that it suggests (sometimes) that the smaller number can be subtracted from the larger only once. Create a more accurate picture that captures the process described.

3.8 (a) Apply the process to numbers 50 and 189. Show all the steps in the process. What do you conclude at the end of the process?

(b) Apply the process to numbers 161 and 70. Show all the steps in the process. What do you conclude at the end of the process ?

3.9 Why does the proposition talk about starting out with two unequal numbers? What will happen if we were to start out with two equal numbers. Discuss.

3.10 In his proof, Euclid says *"Since, then E measures CD, and CD measure BF, therefore E also measures BF"*. Euclid, actually never provides a proof of this statement. Perhaps, it was obvious to him. Provide a proof of the above statement.

3.11 Right after the above statement, Euclid goes on to state:

"But it also measures the whole BA;

therefore it will also measure the remainder AF."

Euclid never provides a proof of this statement. Why should we believe him? Restate in modern notation what Euclid is claiming and provide a proof of this statement.

We now present proposition 2 from Book VII of Euclid's elements. This proposition presents a method to compute the GCD of two numbers which are not prime to each other and provides a proof of the correctness of the method. Euclid's presentation intermixes the proof and the method to some extent. Despite this the elegance of his method and the proof is striking.

PROPOSITION 2.

Given two numbers not prime to one another, to find their greatest common measure.

Let AB, CD be the two given numbers not prime to one another.

Thus it is required to find the greatest common measure of AB, CD.

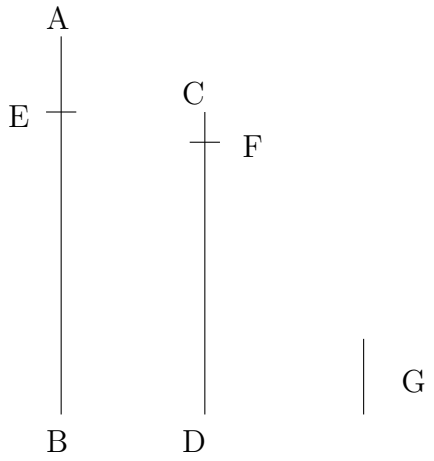
If now CD measures AB - and it also measures itself - CD is a common measure of CD, AB.

And it is manifest that it is also the greatest; for no greater number than CD will measure CD.

But, if CD does not measure AB, then, the less of the numbers AB, CD being continually subtracted from the greater, some number will be left which will measure the one before it.[4]

For an unit will not be left; otherwise AB, CD will be prime to one another [VII, I], which is contrary to the hypothesis.

Therefore some number will be left which will measure the one before it.



Now let CD, measuring BE, leave EA less than itself, let EA, measuring DF, leave FC less than itself, and let CF measure AE.

Since then, CF measures AE, and AE measures DF,

[4]This is the heart of Euclid's description of his algorithm. The statement is somewhat ambiguous and subject to at least two different interpretations. An excercise later in this section explores this issue.

therefore CF will also measure DF.

But it also measures itself;

therefore it will also measure the whole CD.

But CD measures BE;

therefore CF also measures BE.

But it also measures EA;

therefore, it will also measure the whole BA.

But it also measures CD;

therefore CF measures AB, CD.

Therefore CF is a common measure of AB, CD.

I say next that it is also the greatest.

For, if CF is not the greatest common measure of AB, CD, some number which is greater than CF will measure the numbers AB, CD.

Let such a number measure them, and let it be G.

Now, since G measures CD, while CD measures BE, G also measures BE.

But it also measures the whole BA;

therefore it will also measure the remainder AE.

But AE measures DF;

therefore G will also measure DF.

But it will also measure the whole DC;

therefore it will also measure the remainder CF, that is, the greater will measure the less: which is impossible.

Therefore no number which is greater than CF will measure the numbers AB, CD;

therefore CF is the greatest common measure of AB, CD.

Porism. From this it is manifest that, if a number measure two numbers, it will also measure their greatest common measure.

TASKS:

3.12 Read the above proposition, and its proof by Euclid carefully. In particular, Euclid's described method (see fotnote) to compute the GCD is very subtle and is subject to different interpretations. Make sure that you pay attention to all the details to extract the right interpretation.

3.13 Compute the GCD of numbers 1631 and 1008 using your understanding of the method. Show all the steps in your computation.

3.14 The process of repeated subtraction described in proposition 2 and proposition 1 seems identical. Proposition 1 is about numbers prime to one another and proposition 2 is about numbers that are not prime to one another. Why do you think that Euclid wants to separate out these two cases? Is it necessary? Discuss.

3.15 What does Euclid mean when he states *"some number will be left that will measure the one before it."*? In particular, which number does "before" refer to. Explain via an example.

In Euclid's "method" and "proof" the process of continually subtracting the lesser number from the greater is assumed to come to an end within three "swaps" of lesser and greater numbers (CD measuring AE leaves EA less

than itself, EA measuring DF leaves FC less than itself, CF measures AE). For any given pair of numbers this process will be repeated only a finite number of times. However, there is no fixed constant bound on what this finite number might be and depending on what the original numbers are this "finite number of times" can grow in an unbounded fashion. Despite this, Euclid somehow seems to be able to clearly convey what his method is and why it is correct.

3.16 A statement in Euclid's proof goes *"Now let CD, measuring BE, leave EA less than itself.."*. Explain clearly the relationship between the numbers CD, BE and EA in a modern notation and explain how one can compute EA ?

3.17 Explain in your own words Euclid's method of computing the GCD as presented in Proposition 2 of BOOK VII of Euclid.

3.18 Translate your understanding of Euclid's method into "pseudocode" for an algorithm for computing the GCD of two given numbers.

3.19 Prove that the algorithm devised above works correctly to always compute the GCD of two given numbers.

3.20 State the "Porism" (today it will be called a corollary) following Proposition 2 using modern mathematical notation and language and prove that it is indeed correct.

3.21 Implement your algorithm to compute the GCD of two numbers in JAVA/C.

4 Modifying Euclid's Algorithm

You might have noticed that while Euclid's algorithm computes the greatest common divisor of two integers, it never really performs any divisions! It does check for divisibility ("CD measures AB") for purposes of termination. Euclid only uses repeated subtractions in his algorithm. However, Euclid's algorithm can be simplified by using what in computer science lingo is called the ***mod*** operator. The *mod* operator simply returns the remainder upon division: $x \ mod \ y$ is the remainder left when x is divided by y. The remainder is always between 0 and $y - 1$. For example, $7 \ mod \ 3 \ = \ 1$, because 3 can

go into 7 two times (and no more) leaving a remainder of 1. Similarly, 15 mod $5 = 0$ and $3 \; mod \; 7 = 3$.

TASKS:

4.1 Give pseudocode for a ***mod***ified Euclid's Algorithm (pun intended) that eliminates repeated subtraction.

4.2 Implement this modified algoritm using JAVA/C.

4.3 Execute the original and modified algorithms on pairs of large integers and measure the runtime of the two algorithms. Do you encounter any difficulties in the execution of the algorithms? If so, discuss the possible reasons for encountering these problems.

4.4 Do you observe any differences in the runtime of the algorithms? Discuss the possible reasons for observing (or not observing) any differences.

4.5 What would you say are the advantages and disadvantages of the two algorithms?

Comments for the Instructor:

The project is meant for use in an introductory computer science or discrete mathematics course. The project can be used to introduce students to the notion of "computation method" or "algorithm" and to explore concepts like iteration in a basic setting. It allows them to practice their skills in doing proofs but more importantly to observe the evolution of what is accepted as a valid proof or a well-described algorithm. The students will easily notice that the method presented by Euclid to compute the GCD and proof of its correctness that he provided would not be formally accepted as correct today. They will also notice, however, that Euclid is somehow able to convey his ideas behind his method and proof in a way that they can be easily translated into a modern algorithm and proof of its correctness. In this way it will provide them a with sense of connection to the past as well as help them understand that it is possible to develop creative ideas without worrying about formal notation or correctness and that one can use formal notation and proofs to establish the validity of the ideas.

A basic knowledge of programming is essential to successfully complete some of the components of the project.

References

[1] Knuth, D.E., *The Art of Computer Programming*, Volume 1, 1968.

[2] Heath T.L., *Euclid The Thirteen Books of the Elements*, Volume 2, Second Edition, Dover Publications, New York, 1956.