A New Look at Euclid's Second Proposition

Godfried Toussaint

There has been considerable interest during the past 2300 years in comparing different models of geometric computation in terms of their computing power. One of the most well-known results is Mohr's proof in 1672 that all constructions that can be executed with straightedge and compass can be carried out with compass alone. The earliest such proof of the equivalence of models of computation is due to Euclid in his second proposition of Book I of the *Elements*, in which he establishes that the collapsing compass is equivalent in power to the modern compass. Therefore, in the theory of equivalence of models of computation Euclid's second proposition enjoys a singular place. However, the second proposition has received a great deal of criticism over the centuries. Here it is argued that it is Euclid's early Greek commentators and more recent expositors and translators that are at fault, and that Euclid's original algorithm is beyond reproach.

Introduction

In the modern comparative study of geometric algorithms it is imperative first to define the *models of computation*, that is, the characteristics of the machine that will execute the algorithms [30]. A model of computation specifies the number of *processors* used, whether they are used *sequentially* or in *parallel*, the primitive operations allowed, and the cost associated with each of these operations. For example, one of the preferred conceptually abstract models or *ideal* machines in which many geometric algorithms are compared is the *Real RAM* (Random Access Machine) [1], in which each unit of storage space is capable of holding a real number of unlimited precision, and in which the primitive operations that can be performed in one unit of time include the arithmetic operations consisting of addition, subtraction, multiplication, and division, comparisons between real numbers, reading from and writing into a storage location, as well as perhaps some more

Godfried Toussaint

Godfried Toussaint received the Ph.D. degree in Electrical Engineering from the University of British Columbia, Vancouver, B.C., Canada in 1972. Since then he has been Professor of Computer Science at McGill University in Montreal, except for visiting appointments at Stanford, Montreal, Simon Fraser, NYU, University of the West Indies, and Amsterdam. His academic interests include information theory, pattern recognition, and computational geometry. He is on the editorial board of several journals including *Discrete & Computational Geometry*. In 1985, he was awarded a Killam Fellowship by the Canada Council to carry out a two-year research project on movable separability of sets. He enjoys a variety of hobbies, the most passionate being modern dance, African drumming, wind surfing, and scuba diving.

powerful operations, such as computing kth roots, trigonometric functions, and other analytic functions. More controversial assumptions sometimes include the *ceiling* and *floor* functions.

In classical constructive geometry, mathematicians have also been concerned with defining the *models of computation,* that is, the characteristics of the "machine" that will execute the algorithms. Typical machines that have been used in the past starting with Euclid include (1) the *straightedge,* (2) the *ruler,* (3) the *collapsing compass,* (4) the *compass,* (5) the *fixed-aperture compass,* (6) the compass with aperture *bounded from above,* and (7) the compass with aperture *bounded from below,* just to name a few [11, 17, 21, 34]. The *collapsing compass* deserves elaboration here. With the regular compass, one can open it, lock it at a chosen aperture, and lift it off the work space to some other location to draw a circle with the chosen radius. This operation cannot be done with a collapsing compass. The collapsing compass is, like the other machines, an *idealized* machine which allows the compass to be opened to a chosen radius and a circle drawn, but no distance can be *transferred.* It is as if when the compass is lifted off the work space, it collapses and, thus, erases any trace of the previous aperture made.

Of course, more complicated machines can be obtained by combining sets of simple machines. For example, in Euclid's *Elements* he uses the *straightedge* and *collapsing compass* (the combination of machines 1 and 3) as his model of computation. Attempts have also been made to specify the primitive operations allowed with each type of machine [23] and to design constructions that require fewer operations than did Euclid's original constructions. Another active area of research has been to analyze and compare different machine models in terms of their computational power [2, 4, 11, 17]. For example, in 1672 Jorg Mohr [25] and in 1797 the Italian geometer Lorenzo Mascheroni [24] independently proved that any construction that can be carried out with a straightedge and a compass can be carried out with a compass alone; and Jacob Steiner proved in 1833 that the straightedge is equivalent in power to the compass if the former is afforded the use of the compass once [32]. To remind the reader that the straightedge and compass are not yet obsolete computers, we should point out that the Mohr–Mascheroni result was strengthened as recently as in 1987 by Arnon Avron [2] at the University of Tel Aviv.

The earliest proof of the equivalence of models of computation is due to Euclid in his second proposition of Book I of the *Elements,* in which he establishes that the *collapsing compass* is equivalent in power to the *compass.* Therefore, in the theory of equivalence of the power of models of computation, Euclid's second proposition enjoys a singular place. However, like much of Euclid's work and, particularly, his constructions involving many cases, his second proposition has received a great deal of criticism over the centuries. In this article, it is argued that it is Euclid's commentators and translators that are at fault, and that Euclid's original algorithm and proof are beyond reproach.

Euclid's First Two Propositions According to Pedoe

Pedoe [27] contains a lively discussion of Euclid's elements of geometry applied to painting, sculpture, and architecture throughout recent history. To illustrate Euclid's method, he presents the first two propositions of Book 1 of his *Elements.* Earlier in the book Pedoe actually has a completely different algorithm and proof of Proposition 2, to which we shall return at the end of this article. However, at this later point in the book he states that "it is of interest to read how it appears in Euclid." Whereupon the following algorithms and proofs of correctness are presented.

PROPOSITION 1. *On a given finite straight line to construct an equilateral triangle.*
Algorithm 1:
Input: Let AB be the given finite straight line. (Thus, it is required to construct an equilateral triangle on the straight line AB.) See Figure 1.

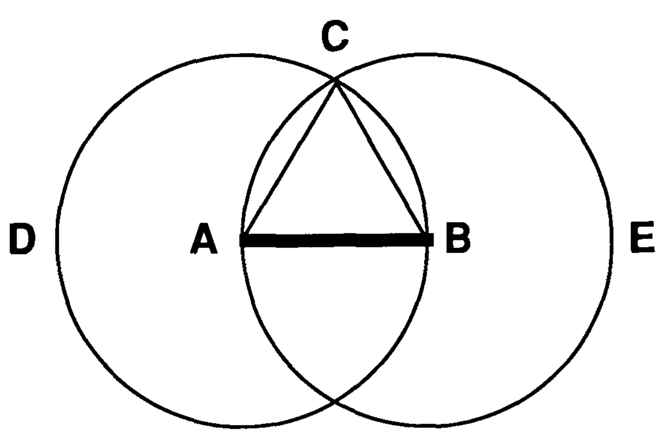


Figure 1. Euclid's figure for the proof of Proposition 1.

Begin
Step 1: With centre A and distance AB let the circle BCD be described.
Step 2: With centre B and distance BA let the circle ACE be described.
Step 3: From the point C, in which the circles cut one another, to the points A, B let the straight lines CA, CB be joined.
End

Proof of Correctness:
Now, since the point A is the centre of the circle CDB, AC is equal to AB. Again, since the point B is the centre of the circle CAE, BC is equal to BA. And things which are equal to the same thing are also equal to one another; therefore, CA is also equal to CB. Therefore, the three straight lines CA, AB, BC are equal to one another. Therefore, the triangle ABC is equilateral; and it has been constructed on the given finite straight line AB. Being what it was required to do.
End of Proof

Of course, neither Euclid nor Pedoe uses the words *algorithm, input, begin,* and *end.* Neither do they use the phrases *proof of correctness* nor *end of proof,* nor do they label separate chunks of the algorithm with the word *Step* such-and-such. However, early Latin manuscripts do preface the construction by the words *exempli causa* and the proof by *probatio eius.* We include these well-known terms found in modern computer science for clarity of layout, and to recall that these divisions did appear in essence in at least the earliest Arab and Latin translations of Euclid's *Elements.* The important thing is that Euclid always gave the algorithm first and the arguments to substantiate the correctness of the algorithm immediately afterward. Even today, too many writers publish geometric algorithms without including a proof of correctness, in spite of the many geometric algorithms that have been found to be incorrect [37]. These authors could certainly take a lesson here from Euclid. Sometimes, as we shall see below, the algorithms in the *Elements* include unnecessary steps for obtaining the solution, but these steps have the benefit of simplifying the ensuing proof of correctness.

Euclid also made use of another common practice in modern computer science: *subroutines.* In the algorithm of his second proposition described next, he uses **Algorithm 1** above. Below we give Pedoe's description of Euclid's construction.

PROPOSITION 2. *To place at a given point (as an extremity) a straight line equal to a given straight line.*
Algorithm P [Pedoe's version]:
Input: Let A be the given point, and BC the given straight line. (Thus, it is required to place at the point A (as an extremity) a straight line equal to the given straight line BC.) See Figure 2.
Begin
Step 1: From the point A to the point B let the straight line AB be joined.
Step 2: On AB (using **Algorithm 1**) let the equilateral triangle DAB be constructed.
Step 3: With centre B and distance BC let the circle CGH be described.
Step 4: With centre D and distance DG let the circle GKL be described.

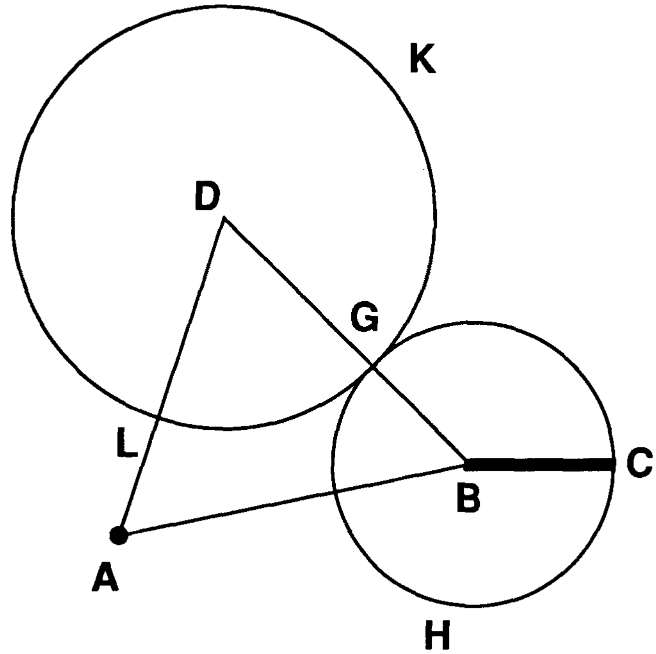


Figure 2. Pedoe's figure for proving Euclid's Proposition 2.

Exit with AL as the solution.
End
Proof of Correctness:
Then, since the point B is the centre of the circle CGH, BC is equal to BG. Again, since the point D is the centre of the circle GKL, DL is equal to DG, and in these DA is equal to DB. Therefore, the remainder AL is equal to the remainder BG. But BC was also proved equal to BG. Therefore, each of the straight lines AL, BC is equal to BG; and things which are equal to the same thing are also equal to one another. Therefore, AL is also equal to BC. Therefore, at the given point A the straight line AL is placed equal to the given straight line BC. Being what it was required to do.
End of Proof

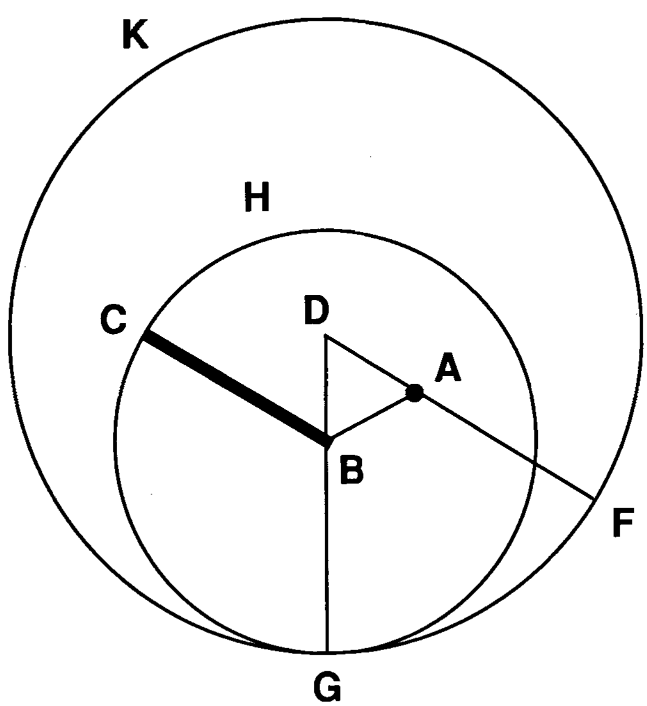
We remark here that Pedoe's figure, shown in Figure 2, is considerably different from those in other sources on Euclid such as Heiberg [16], Heath [15], and Dijksterhuis [13] for example. Much more serious, however, is the fact that **Algorithm P** given by Pedoe is incorrect! It is clear that for a solution to be obtained by **Algorithm P,** it is crucial that the circle centred at B with radius BC intersect DB at G. Otherwise, G is undefined and the rest of the algorithm makes no sense. Now consider what happens when the length of BC is greater than the distance from A to B. Clearly, the circle centred at B with radius BC will completely enclose triangle ABD and its interior and the construction fails! In modern parlance, for such an input the algorithm crashes.

Euclid's Construction According to 19th-, 18th-, and 17th-Century Scholars

During the 19th century, along with more than 700 editions of *The Elements*, there was a flurry of textbooks on Euclid's *Elements* for use in the schools and colleges. A sample of several of these books [14, 22, 33, 36] yields a common (apart from notation) algorithm and illustrative figure for Euclid's second proposition. However, both the algorithm and figure are quite different from Pedoe's. Consider then the algorithm according to one of these sources [14].

PROPOSITION 2. *From a given point to draw a straight line equal to a given straight line.*
Algorithm 19C [Popular 19th-century version]:
Input: Let *A* be the given point and *BC* the given straight line. (It is required to draw from the point *A* a straight line equal to *BC*.) See Figure 3.
Begin
Step 1: Join *AB*.
Step 2: On *AB*, describe an equilateral triangle *DAB*.
Step 3: From centre *B*, with radius *BC*, describe the circle *CGH*.
Step 4: Produce *DB* to meet the circle *CGH* at *G*.
Step 5: From centre *D*, with radius *DG*, describe the circle *GKF*.
Step 6: Produce *DA* to meet the circle *GKF* at *F*. Then *AF* shall be equal to *BC*.
End

Figure 3. Popular 19th-century figure for the proof of Euclid's Proposition 2.

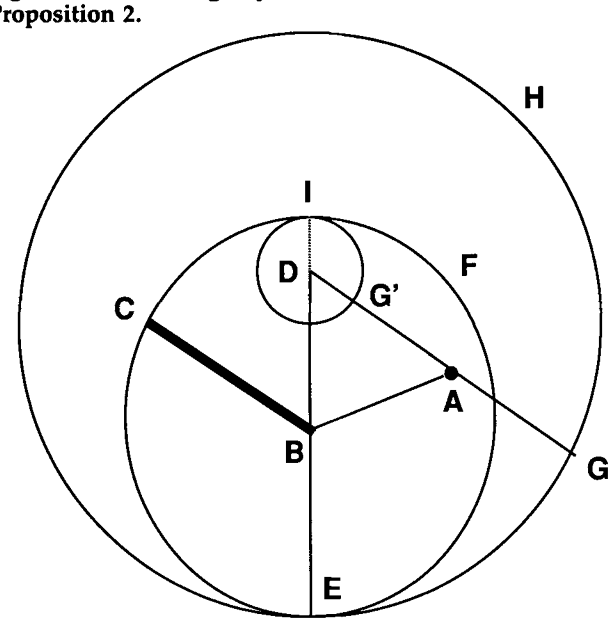


This algorithm is certainly an improvement over Pedoe's algorithm as it appears to work correctly for some input configurations whether *BC* is greater than or less than *BA*. Nevertheless the algorithm suffers from ambiguous statements. Step 4 asks us to produce (extend in length) *DB* to meet the circle *CGH* at *G*, but it does not tell us in which direction (emerging from *D* or from *B*) to produce *DB*, and certainly in either direction we are bound to meet the corresponding circle constructed in Step 3. Figure 3 shows one possible case, but had we produced *DB* in the direction from *B* to *D* instead of the direction shown we would have obtained a completely different intersection point *G*. A similar problem exists with Step 6.

The ambiguities observed in the algorithms described in [14] and [33], which are exemplified here as **Algorithm 19C,** are absent in the exposition by Taylor [35], if not in the body of the algorithm at least in the subsequent discussion, where it is indicated that we are free to choose one or the other alternative as in Step 1. It is therefore instructive to examine his algorithm and accompanying discussion in more detail.

PROPOSITION 2. *From a given point to draw a straight line equal to a given straight line.*
Algorithm T [Taylor's version]:
Input: Let *A* be the given point and *BC* the given straight line. (It is required to draw from the point *A* a straight line equal to *BC*.) Refer to Figure 4.
Begin
Step 1: Draw *AB*, the straight line from *A* to one of the extremities of *BC*.
Step 2: On it, construct an equilateral triangle *DAB*.
Step 3: With *B* as centre and *BC* as radius, describe the circle *CEF*, meeting *DB* (produced if necessary) at *E*.

Figure 4. Illustrating Taylor's version of Euclid's Proposition 2.



Step 4: With D as centre and DE as radius, describe the circle EGH, meeting DA (produced **if** necessary) at G. Then AG is the straight line as required.
End

Note that Taylor is careful to add in Steps 3 and 4 the explicit *if statements* that DB and DA are to be produced if necessary. Therefore we presume that if the construction circle CEF intersects the sides of equilateral triangle ABD, then the extension of DA need not be carried out. Unlike the previous 19th-century geometry books, Taylor follows the proof of Proposition 2 with the following interesting discussion.

> It is assumed in this proposition that the straight line DB intersects the circle CEF. It is easily seen that it must intersect in two places.
>
> It will be noticed that in the construction of this proposition there are several steps at which a choice of two alternatives is afforded: (1) we can draw either AB or AC as the straight line on which to construct an equilateral triangle: (2) we can construct an equilateral triangle on either side of AB: (3) if DB cut the circle in E and I, we can choose either DE or DI as the radius of the circle which we describe with D as centre.
>
> There are therefore three steps in the construction, at each of which there is a choice of two alternatives: the total number of solutions of the problem is therefore $2 \times 2 \times 2$ or eight.

We see that Taylor's way of dealing with the ambiguities discussed above is to explicitly acknowledge that there are eight different cases to Euclid's proposition that depend on how the construction is carried out, that we are free to choose any one of these eight paths through the implied decision tree, and that the sides DB and DA need not be produced if not necessary. In light of this classification, let us follow down one path of these choices on the input configuration illustrated in Figure 4, where it is assumed that the length of CB is greater than the length of CA. In our first choice we, therefore, select AB as the segment on which to construct our equilateral triangle. Our second decision will be to construct the triangle on the side shown in Figure 4. Now because the circle CEF does not intersect the triangle, we extend DB, which cuts the circle at the two points E and I. According to Taylor, we may now choose either DE or DI as the radius of the circle which we describe with D as centre. Let us choose DI. Now, this circle with D as centre intersects DA at G', playing the role of G in his algorithm, and, therefore, according to Step 4, DA need not be produced to G. According to the algorithm, therefore, the solution is given by AG' which is clearly incorrect, because AG' is smaller than AB, whereas BC is greater than AB, by assumption. Therefore, although the ambiguities of **Algorithm 19C** have been removed by Taylor, **Algorithm T** does not always yield the correct solution on a given line–point configuration, depending on which construction strategy is applied.

Furthermore, **Algorithm T** suffers from an additional minor bug not even present in **Algorithm 19C.** Note that Step 1 in **Algorithm 19C** does not offer choice. **Algorithm T** asks that A be connected to one of the extremities of BC, one that we are free to choose. However, if we choose to connect A to C (rather than B as in Taylor's figure), then it is impossible to execute Step 2, and the algorithm crashes.

Another author, Lardner [22], also follows his presentation of an ambiguous algorithm identical to **Algorithm T** with a discussion of how the student should be careful about different cases arising from the varieties of different input configurations. In his own words,

> The different positions which the given right line and the given point may have with respect to each other, are apt to occasion such changes in the diagram as to lead the student into error in the execution of the construction for the solution of this problem.
>
> Hence it is necessary that in solving this problem the student should be guided by certain *general* directions, which are independent of any particular arrangement which the several lines concerned in the solution may assume. If the student is governed by the following general directions, no change which the diagram can undergo will mislead him.

Lardner then proceeds to present six general rules concerning what can and cannot be done to ensure that **Algorithm T** works correctly on all inputs. This discussion includes a case analysis of construction strategies. Unlike Taylor [35], it does not allow DA and DB to be extended in either direction, but insists that they be extended through the base of the constructed triangle, thus concluding that the solution to Euclid's second proposition has four cases rather than Taylor's eight. Another general rule that Lardner insists should be followed is that the centre of the circle constructed in Step 3 should lie at the extremity of BC connected to A in Step 1, thus avoiding one of Taylor's problems.

Another variation occurs in a much earlier Scottish book on Euclidean geometry published in 1831 by John Playfair [29], which has a variant of **Algorithm 19C.** In this book, we are asked to extend DA and DB to E and F, respectively, and thus the ambiguity of **Algorithm 19C** is also present here. However, unlike **Algorithm 19C** or **Algorithm T,** the algorithm in [29] first performs the extensions and subsequently constructs the circles.

We close this section with a note on textbooks of the 18th and 17th centuries. In these two centuries combined the number of editions of Euclid's *Elements* published was less than half of the number for the 19th century, about 325 and 280 in the 18th and 17th centuries, respectively. It is also much more difficult to find copies of these earlier editions. I have held in my hands only two editions from the 18th century [5, 38] and one from the 17th century [10], having found all three in the special collection of the library at Queens University in Kingston, Ontario. The 1705 manuscript by Isaac Barrow (from Trinity College, Cambridge) has the additional distinction (according to the claim on its front page) of being the first English copy translated from Latin. What is worth noting about the algorithms in these texts is that (1) they are identical to each other; (2) *like* **Algorithm 19C,** they are ambiguous; but (3) *unlike* all other algorithms I have encountered, they begin not by connecting point A to one of the end points of segment BC but by constructing a circle of radius BC centred at one of the end points of BC. Then, in the second step, point A is joined to the end point selected as the centre in the previous step. Note that this ordering circumvents the problem that **Algorithm T** has with Steps 1 and 2, and furthermore allows us to ignore Lardner's caveat intended to resolve it.

Euclid's Construction According to Gerard of Cremona and Peyrard

One is naturally led to the question: Which of all these algorithms is the one Euclid originally proposed? It would be easy to answer this question by looking up Euclid's original manuscript. Unfortunately, history has made this impossible. In the year 332 B.C. Alexander the Great, at the age of 24, conquered Egypt and founded the city of Alexandria. When, after conquering much of the rest of the world, he died at the age of 33, his generals divided up the empire. In this way Egypt fell into the hands of Ptolemy I in 306 B.C. Ptolemy II created the University of Alexandria, which became, by virtue of its excellent scholars (including Euclid) and its impressive library (three-quarters of a million books including Euclid's original version of *The Elements*), the intellectual and scientific centre of the world. In 48 B.C. Julius Caesar occupied Alexandria and intended to carry a large portion of the library with him back to Rome. The academic community held a demonstration which was quickly quelled by Caesar's army. In the fighting many of the books were burned. More books were burned during later Egyptian revolts in 272 A.D. and 295 A.D. In the 4th and 5th centuries, zealous Christian bishops began to persecute the pagan writers (mathematicians) and their books. Bishop Theophilus in 391 A.D. led a Christian mob and destroyed the Temple of Serapis which housed many of the remaining books. The last mathematician alive in Alexandria, a woman by the name of Hypatia, was hacked to pieces by Bishop Cyril. Finally, the Arabs invaded Egypt in 646 A.D., and General Amr ibn-al-As burned the remaining books, allegedly because [6] "if the books agreed with the Koran they were superfluous; if they disagreed they were pernicious." In short, in all likelihood Euclid's original algorithm went up in smoke.

In spite of the criticism often directed at Euclid, one may find it difficult to believe that he could have been guilty of such oversight as in the versions of his algorithm exhibited so far. However, established authorities on Euclid, such as Heiberg [16], Heath [15], and Dijksterhuis [13], have a significantly different algorithm. The figure in these three works is given in Figure 5, and the algorithm is given below. We omit the proof of correctness as it is exactly the same as that given by Pedoe.

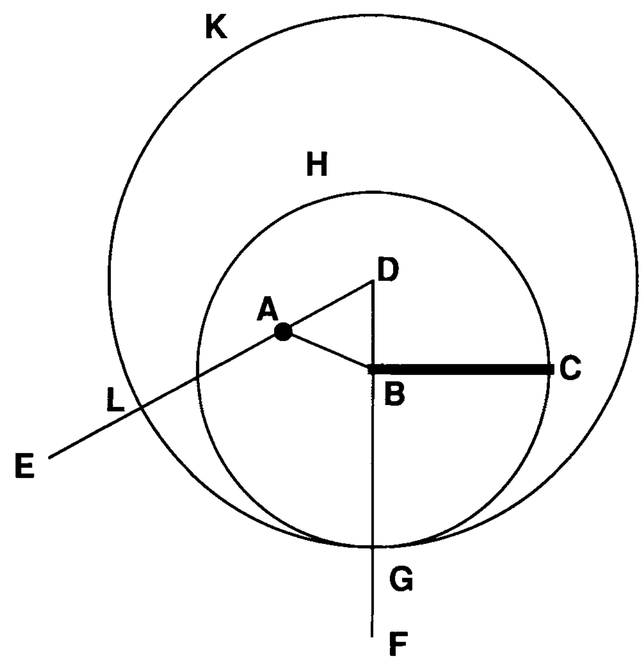


Figure 5. Euclid's figure for the proof of Proposition 2 according to Heath.

PROPOSITION 2. *To place at a given point (as an extremity) a straight line equal to a given straight line.*
Algorithm Euclid [Heath's version]:
Input: Let A be the given point, and BC the given straight line. (Thus, it is required to place at the point A (as an extremity) a straight line equal to the given straight line BC.) See Figure 5.
Begin
Step 1: From the point A to the point B let the straight line AB be joined.
Step 2: On AB (using **Algorithm 1**) let the equilateral triangle DAB be constructed.
Step 3: Let the straight lines AE, BF be produced in a straight line with DA, DB.
Step 4: With centre B and distance BC let the circle CGH be described.
Step 5: With centre D and distance DG let the circle GKL be described.
Exit with AL as the solution.
End

Note that in Figure 5 the length of *BC* is indeed larger than the distance between *A* and *B* and Pedoe's version of Euclid's algorithm would not work in this case. However, for unknown reasons (I will offer a conjecture later), Pedoe leaves out Step 3 in the above algorithm. This crucial step in Euclid's algorithm constructs the extensions of *DA* and *DB* in directions *E* and *F*, respectively, thus ensuring that, whether or not the length of *BC* is larger than the distance between *A* and *B*, the algorithm continues to "execute," and the figure remains the same in the sense that point *G* exists and lies on *BF*. Note the difference between the manner in which *DA* and *DB* are to be produced in **Algorithm Euclid** as compared to **Algorithm 19C** and **Algorithm T.** In the latter two algorithms, the ambiguous instructions state that the *sides of the equilateral triangle DA and DB are to be produced.* In **Algorithm Euclid,** on the other hand, the statement in Step 3 concerning the extension of *DA* and *DB* states, "Let *the straight lines AE, BF be produced in a straight line with DA, DB.*" In other words, the extensions are to be collinear with (in a straight line with or in the direction of) *DA* and *DB*. No room is left here for choosing the direction of the extensions of *DA* and *DB* as in **Algorithm 19C.**

At this point, one may wonder about the authenticity and correctness of the accounts of Heiberg [16], Heath [15], and Dijksterhuis [13]. The Greek text by Heiberg is considered to be the definitive edition. It consists of portions taken from different Greek manuscripts spanning the 9th to 12th centuries and considered by philologists to be the most authentic. There also exist several interesting Latin manuscripts which are translations of Arabic manuscripts. We may peruse the first printed edition of the Latin translation of the Arabic (Ishaq-Thabit) version of Euclid's *Elements,* believed to have been made by the monk Gerard of Cremona in Toledo during the 12th century [8] following its discovery in Baghdad.

We find that apart from the letters E and F in Heath being replaced in [8] by L and G, respectively, the algorithms and proofs of correctness found in [13, 15, 16] are identical to those in the 12th-century manuscript. This 12th-century algorithm is a Latin translation of an Arabic translation of a Theonian Greek manuscript. In fact, all Arabic translations are believed to descend from the recension by Theon of Alexandria.

Anyone who has played the translation game may wonder how this version compares with early Greek manuscripts with respect to the crucial Step 3. In another 12th-century Sicilian Latin translation (of unknown authorship) of Euclid's *Elements* made directly from the Greek [9], Step 3 is stated as follows:

> Educantur *in directo rectis DA et DB recte AE et BF.*

This translates to "Lead *forth the straight lines AE and BF in a straight line with (in the direction of) the straight lines DA and DB*" and is, thus, in agreement with the Gerard of Cremona version and Heiberg's definitive edition.

A final piece of evidence that **Algorithm Euclid** described above is indeed Euclid's is the so-called manuscript P, the Vatican manuscript No. 190. Until 1804, all manuscripts of Euclid's *Elements* were believed to be descended from Theon's 4th-century recension [7]. When Napoleon conquered Italy, he stole from the Vatican a Greek manuscript (No. 190) of Euclid's *Elements* which he kept in the King's Library in Paris. F. Peyrard, a professor at the Lycée Bonaparte, wanted to write a definitive French version of the *Elements* using the best Greek manuscripts at his disposal, and towards that end obtained access to the King's Library. There he found manuscript No. 190, and to his astonishment discovered he had in his hands a pre-Theonian 10th-century manuscript. In the meantime the Allied Forces defeated Napoleon and forced France to return all stolen works of art. At the request of the French government, the Pope made Peyrard a happy man by granting an extension of the return date of the manuscript, thus giving him enough time to finish his translation [28]. In Peyrard's manuscript, which he emphasizes is a literal translation, the crucial Step 3 is written as "Menons *les droites AE, BZ dans la direction de DA, DB,*" and is, thus, in agreement with **Algorithm Euclid** described above.

Cases in Constructions and Proofs

The above discussion brings up naturally the general question of the analysis of *cases* in Euclid's constructions, modern computational geometry, and geometric proofs in general. When we talk about *cases* today we generally mean equivalence classes of input configurations, rather than instances of the construction sequence resulting from a set of choices made as a result of the ambiguities of the algorithm's description, as are the case classifications in Taylor [35] and Lardner [22]. An algorithm must be specified unambiguously and should execute correctly for all inputs it was designed to handle.

Much criticism has been heaped on Euclid over the past two thousand years for his alleged sloppiness in constructions and proofs concerning cases. For one thing, he has been accused of giving proofs of correctness that depend severely on the figure accompanying the proof. Thus, Bertrand Russell [31]:

> A valid proof retains its demonstrative force when no figure is drawn, but very many of Euclid's earlier proofs fail before this test . . . The value of his work as a masterpiece of logic has been very grossly exaggerated.

Again, in the words of William Dunham [12]:

> Admittedly, when he allowed himself to be led by the diagram and not the logic behind it, Euclid committed what we might call a sin of omission. Yet nowhere in all 465 propositions did he fall into a sin of commission. None of his 465 theorems is false.

Finally, in the words of Felix Klein [20]:

> Euclid . . . always had to consider different cases with the aid of figures. Since he placed so little importance upon correct geometric drawing, there is real danger that a pupil of Euclid may, because of a falsely drawn figure, come to a *false* conclusion.

A proposition that has a plethora of cases and that has been the subject of much criticism of Euclid is in fact Proposition 2, the topic of this article. It will be argued here using this proposition as a "case" study that much of the criticism of Euclid's case analysis stems from a lack of understanding of his original work due in part to the writings of the early Greek commentators of the *Elements* such as Heron and Theon of Alexandria and others reviewed by Proclus [26] in the 5th century, and exacerbated by a 12th-century Latin translation of an Arabic manuscript by Adelard of Bath [7] and many English scholars of the 19th century. Furthermore, if we judge the original algorithm and proof of correctness of Euclid's *Proposition 2* using today's highest standards in the field of computational geometry Euclid deserves praise for his brilliance.

Return then to Euclid's second proposition: *To place at a given point (as an extremity) a straight line equal to a given straight line.* Clearly, an algorithm for carrying out this task has to execute, that is, be well defined for all inputs, that is, for all possible line segments *BC* and all points *A* no matter how they are positioned with respect to each other in the plane. Furthermore, unless the algorithm is designed to work only for inputs in *general position,* it should also be able to handle singularities such as when point *A* lies on the segment *BC* or *A* is equidistant from *B* and *C*. Similarly, a proof of correctness must establish that in all situations the algorithm will yield the correct solution.

Euclid had the habit, as is well illustrated by Figure 5, of including only one figure to illustrate the construction and proof. A reader may thus wonder, on stepping through the algorithm on the given figure, whether the same steps would work on a completely different figure. The same reader may even be skeptical as to whether the arguments in the proof of correctness would carry over with the same letters used as labels of crucial points derived during the construction. This, in fact, appears to have been the reaction of early Greek commentators of the *Elements,* who criticized Euclid for leaving out cases that they discovered missing and then supplied proofs of their own. An in-depth commentary of Euclid's elements and subsequent criticisms made against it was written in the 5th century by Proclus [26]. Proclus himself does not usually criticize Euclid, and on several occasions actually comes to his defense. In the words of Glenn Morrow [26],

> When in the proof of a theorem Euclid uses only one of two or more possible cases, as is his custom, Proclus will often prove one or more of the omitted cases; sometimes he simply calls attention to them and recommends that his readers, "for the sake of practice," prove them for themselves. Sometimes he gives an alternative proof of a theorem devised by one of his predecessors for the obvious purposes of showing the superior elegance or appropriateness of Euclid's demonstration.

Indeed one can conjure up many special cases of an initial configuration of point *A* and line segment *BC*. For example: *Case 1: A* may lie on the line collinear with *BC*, or *Case 2: A* may lie on one side of the line collinear with *BC*. In Case 1, *A* may lie on the line segment *BC* (Case 1.1) or off the line segment *BC* (Case 1.2). If *A* lies on *BC*, then in Case 1.1.1 it may lie on an end point of *BC* or in Case 1.1.2 on the interior of *BC*, and in the latter case we have two more cases depending on whether *A* is closer to *B* or closer to *C*. In Case 1.2, where *A* lies off segment *BC*, *A* could be closer to *B* (Case 1.2.1) or to *C* (Case 1.2.2). Furthermore, Case 1.2.1 divides into two more cases depending on whether the distance between *A* and *B* is greater than or less than the distance between *B* and *C*. Case 2 in which *A* lies off the line collinear with *BC* can also be divided into cases using a variety of criteria. For example, we might consider two cases depending on whether the line segment *BC* lies in the interior (Case 2.1) or the exterior (Case 2.2) angle that triangle *ABD* makes at *D*. Finally, each of these two cases determines two more cases depending on whether the distance between *A* and *B* is greater than or less than the distance between *B* and *C*.

Some of the above cases (but certainly not all!) were discussed by the Greek commentators and are included in the work of Proclus. Usually a proof that Euclid's algorithm worked correctly was then provided for the particular case at hand. Sometimes the actual *algorithm* given by Euclid was changed to handle the special case. For example, for a particular input configuration in Case 2.1 with the distance between *A* and *B* less than the distance between *B* and *C*, Proclus objects to Euclid's algorithm because line segment *BC* "gets in the way" of the construction of triangle *ABD* above segment *AB* (see Fig. 6). In the words of Proclus, "for there is not room." Heath notes that Heron of Alexandria circa 100 A.D., in his commentary on the *Elements,* also sometimes used constructions different from Euclid's to circumvent objections of this type. The algorithm of Proclus for this particular case follows (see Fig. 6).

Begin
Step 1: Let a circle be drawn with centre at *B* and distance *BC*.
Step 2: Let the lines *AD* and *BD* be produced to *F* and *G*.
Step 3: With centre at *D* and distance *DG* let the circle *GE* be described.
[Exit with *AE* as the solution.]
End

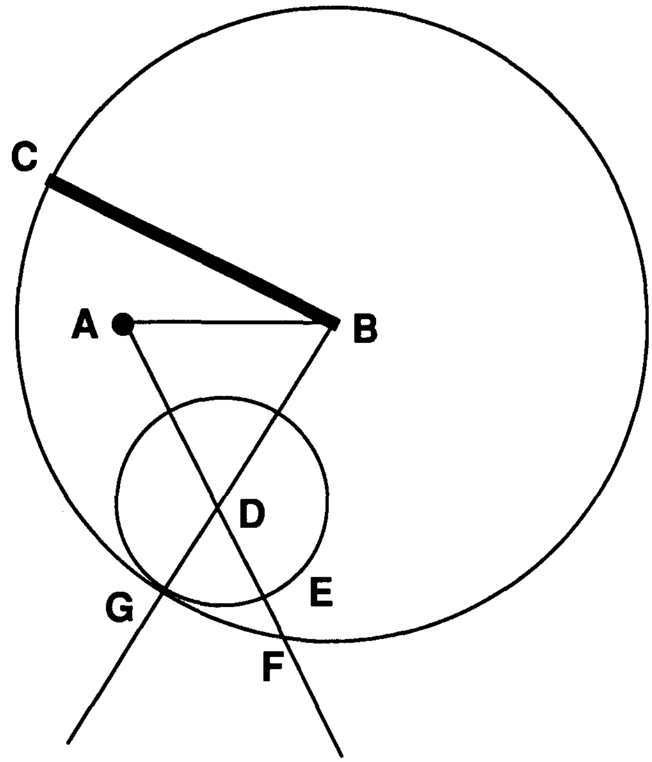


Figure 6. Proclus's figure for the proof of a subcase of Case 2.1 of Proposition 2.

Note how Proclus has changed the clear line-extension statements of Euclid's algorithm to the ambiguous statements (let the lines AD and BD be produced to F and G) found in the 19th-century accounts, and that the correctness of the construction is made to depend on the figure!

Another fascinating manuscript is an Arabic book titled *On the Resolution of Doubts in Euclid's Elements and Interpretation of Its Special Meanings* written in 1041 A.D. by Ibn al-Haytham. A copy of this book made in 1084 A.D. was found in the University of Istanbul Library [18]. As the title suggests, this is not a translation of the *Elements* but a discussion about well-known criticisms of Euclid's work. In discussing Euclid's second proposition al-Haytham treats four basic cases in terms of input: (1) point A is either B or C, (2) A lies on the line segment BC, (3) A lies on the line passing through BC, and (4) A lies outside the line passing through BC. In addition to these, he has a very strange case that does not appear to have been mentioned anywhere else, and this is the case when the line segment BC and the point A are separated by a valley or a river so that the line joining the points A and B cannot be drawn! His solution to this last case is most puzzling, for he writes that the way to handle this case is to measure the line segment and redraw it in the neighborhood of the point, after which Euclid's procedure is then applied! It would appear that Ibn al-Haytham was not lacking a sense of humor in his mathematical writings.

Euclid's Algorithm Reconsidered

It is clear from the above discussion that Euclid's followers were concerned that perhaps Euclid's algorithm and proof of correctness did not hold for all possible configurations of the input to the problem. I will argue that the commentators themselves succumbed to the fallacy of "going by the figure" even more than Euclid himself, and that they missed the *essence, semantics,* or *deep structure* behind Euclid's proof.

First, we should remember that when cases did in fact exist, Euclid used figures to *illustrate* a construction and proof rather than make a *case* statement. In the words of Heath,

> To distinguish a number of cases in this way was foreign to the really classical manner. Thus, as we shall see, Euclid's method is to give one case only, for choice the most difficult, leaving the reader to supply the rest for himself. Where there was a real distinction between cases, sufficient to necessitate a substantial difference in the proof, the practice was to give separate *enunciations* and proofs altogether.

This is, indeed, the social convention followed even today in computational geometry, where the phrase "the remaining cases can be proved in a similar way" is seen in almost every published paper in the most scholarly of journals.

I conjecture, though, that Euclid saw no cases in Proposition 2 because fundamentally there are not any. Furthermore, if the reader will follow through Euclid's original algorithm in all the possible "fabricated" cases enumerated in the previous section, he or she will find that the algorithm is well defined in the modern sense and will execute correctly and terminate with the correct solution. Furthermore, the proof of correctness also follows through. This cannot be said of any of the subsequent algorithms and proofs offered by Heron, Proclus, and the other Greek commentators of Euclid, nor the 19th-century English scholars. It should be mentioned here that one logical (out-of-context) situation consists of Case 1.1.1 in which the point A lies at one end point of segment BC. Clearly, in this pathological situation an equilateral triangle cannot be constructed on AB and the algorithm would be undefined and fail to execute. However, the purpose of the problem is *to transfer a distance.* If A coincides with either B or C, then the answer, namely, segment BC, is already known at the start. Therefore, the algorithm is clearly intended to work for all points A on the plane except B and C.

The reader may experience an interesting effect upon actually carrying out Euclid's construction and proof for all the cases enumerated above, and that is the *Eureka* experience in which the *essence, semantics,* or *deep structure* behind Euclid's construction is made manifest. Once this happens, it is transparently clear that Euclid's algorithm and proof of correctness are valid for all cases one could possibly imagine. Fundamentally there are indeed no cases.

It is difficult to grasp the essence of the algorithm-proof by fixing an input configuration and then analyzing variations in constructions as in the work of the Greek commentators. However, the following "trick" makes the essence "jump out of the page at you." We fix the construction instead and for this fixed construction we "look" at all possible input configurations. The crucial part of Euclid's construction (missing in Pedoe's algorithm [27] and missed by most of Euclid's followers) is the *cone* determined by the rays *DE* and *DF* and making an angle of 60° at *D*. This cone is implicitly constructed by the resulting concatenation of the equilateral triangle *DAB* and the extensions constructed in Step 3 from *A* to *F* and *B* to *E*. This cone is as large as desired and always subtends 60°. Consider such a cone as fixed in space and refer to Figure 7. Now point *A* must always lie on one ray *DF*. Also, line segment *BC* must always have its end point *B* on the other ray *DE*. With the compass anchored on *B*, Euclid's construction first marks off a point *G* on *BE* such that *BG* equals *BC*. Then with the compass anchored on *D*, it marks off a point *L* on *AF* such that *DL* equals *DG*. It is clear that for all possible configurations of points *A* and line segments *BC* the construction is valid. Variation in the distance between *A* and *B* does not change the essence of the proof. Furthermore, all possible relative positions of segment *BC* with respect to point *A* retain their property of cutting *BE* at *G*. It does not matter whether *BC* is greater than, less than, or equal to *AB*. Neither does it matter if *C* lies on *AB* or *DB*, or for that matter if it coincides with point *A* or *D*! Therefore, the algorithm is well defined and executes in all possible cases.

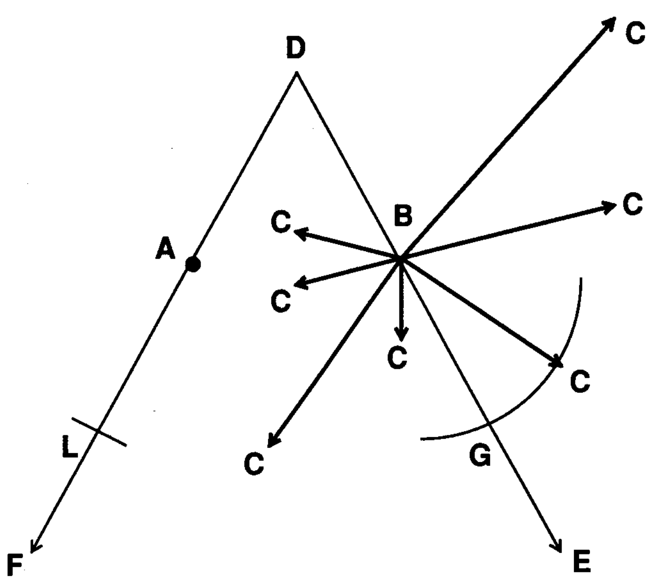


Figure 7. Illustrating the proof of Euclid's Proposition 2 for all cases.

Because in all cases *DB* equals *DA*, it follows that the algorithm yields the correct solution in all cases as well.

This then is the *logic* behind Euclid's proof; and, we might add that, Bertrand Russell [31] and Dunham [12] notwithstanding, it holds without the need of a figure. We see at once Euclid's brilliance in the extension of *DA* and *DB* in the directions of *A* and *B* to create the *cone* with apex at *D* rather than in the direction of *D* as done by Proclus, for example. It is also easy to see with the aid of this cone that indeed there are no proper cases here at all. The cases fabricated and considered by Euclid's commentators are artifacts of a lack of understanding of the underlying logic which, it is conjectured, Euclid had in mind. In light of the culturally established belief held by so many that Euclid's proofs only hold for certain cases, together with the fact that almost all modern versions of the construction are either ambiguous or downright incorrect, it is easy to understand why Pedoe [27] picks only one such case and claims to give Euclid's original proof although it is missing the crucial construction of the cone mentioned above.

We close this section with a conjecture as to how it came about that so many of the English textbooks contain an incorrect algorithm for Euclid's second proposition. I believe the answer may lie in the famous Latin translation (of an Arabic manuscript by Al-Hajjaj) due to Adelard of Bath [7].

Among the most well-known medieval English translators of Euclid's *Elements* was Adelard of Bath in the 12th century. Actually, Adelard of Bath's name is associated with three distinct versions of the *Elements;* according to Busard [7], it was version II "that became the most popular of the various translations of the *Elements* produced in the 12th Century and apparently the one most commonly studied in the schools." Furthermore, this version is apparently the least authentic. In the words of Busard, "not only are the enunciations differently expressed but the proofs are very often replaced by instructions for proofs or outlines of proofs."

Adelard of Bath writes Step 3 as follows:

Protrahanturque linee DA et DB directe usque ad L et G.

His actual letters are different and are here substituted to match those of Figure 5 for ease of discussion. As a minor aside, there is an error (probably typographic?) in this manuscript, that is, *L* and *G* are actually reversed. More seriously, *E* and *F* are nonexistent, as are the references to producing the lines *AF* and *BE*, and the literal translation reads "Draw forward (extend) lines *DA* and *DB* until *L* and *G*." The sentence that pervades the English textbooks reads, "Produce lines *DA* and *DB* until *L* and *G*." Thus, one possibility is that Adelard of Bath is responsible for introducing the error. To be sure, it is known that in the 4th-century Theon of Alexandria's recension of the *Elements* involved altering the language in some places and sometimes supplying alternative proofs; and according to Busard [7] all the manuscripts of the *Elements* known until the 19th century were derived from Theon's text. Is Theon the culprit here? Adelard of Bath translated his manuscript from the Al-Hajjaj manuscript in Arabic; one may wonder if Al-Hajjaj is to blame. However, it is generally considered that the Arabic manuscripts are quite trustworthy, and other Latin translations of Arabic manuscripts, such as that of Gerard of Cremona, have a correct algorithm. The finger seems to point in the direction of Adelard of Bath.

20th-Century Algorithms

For the sake of comparison, contrast, and completeness, we offer in this section an alternative modern construction that is fundamentally different from all those considered by Euclid, Heron, Proclus, and the other Greek and subsequent commentators, as well as the plethora of 19th- and early 20th-century textbook writers. It is based on the notion of *mirror symmetry.*

Recall that in 1672 Jorg Mohr and in 1797 the Italian geometer Lorenzo Mascheroni independently proved that any construction that can be carried out with a straightedge and a compass can be carried out with a compass alone. The reader may wonder how on earth we can draw a line segment of length BC with one extremity at A *without* using a straightedge. Strictly speaking, we cannot, and therefore in constructions of a line or line segment with compasses alone, we require only that two points on the line or the two end points of the line segment be specified. Thus, we are actually *simulating* a line or line segment by two points. In this sense it is more appropriate to state the Mohr–Mascheroni theorem as: *any construction that can be carried out with a straightedge and a compass can be* simulated *with a compass alone.* The above constructions use both a straightedge and a compass. It is fitting to end this discussion with a construction that uses a compass only. We present the one described in [17] which is also the first construction presented in Pedoe [27].

Algorithm CO [Compass Only version]:
Input: Let A be the given point and BC the given straight line. [Thus, it is required to place at the point A (as an extremity) a straight line equal to the given straight line BC.] See Figure 8.
Begin
Step 1: Draw a circle with centre A and radius AB.
Step 2: Draw a circle with centre B and radius BA. (The two circles intersect at D and E.)
Step 3: Draw a circle with centre D and radius DC.
Step 4: Draw a circle with centre E and radius EC.

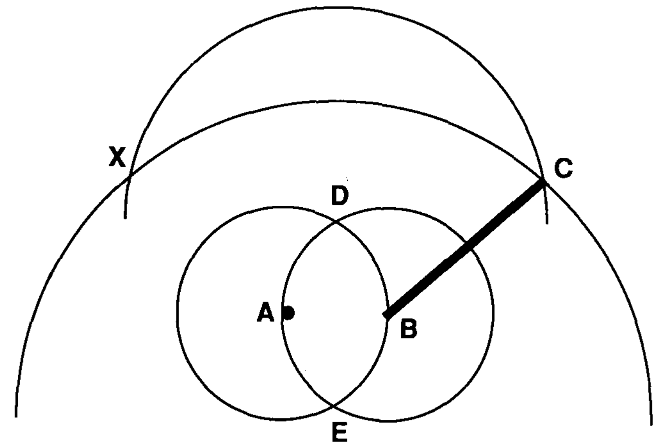


Figure 8. Illustrating the construction with compasses only.

These two circles intersect at C and X where X is the desired reflection point of C across the imaginary line through DE, and XA is the desired length.
End

In the spirit of Proclus we invite the reader to supply the proof of correctness.

Conclusions

We mention in closing that even the 20th-century **Algorithm CO** pales by comparison with **Algorithm Euclid** from the point of view of robustness with respect to singularities. Consider, for example, the case where point C happens to lie at a location equidistant from A and B. **Algorithm Euclid** executes in this case as easily as in any other because everything is well-defined. Without special flag-waving code, however, **Algorithm CO** could crash attempting to draw a circle with radius zero and then intersecting two circles, one of which has radius zero.

One apparent difference between modern and classical computational geometry concerns the issue of lower bounds on the complexity of geometric problems. Although Lemoine [23] and others were concerned with defining primitive operations and counting the number of such operations involved in a construction, they do not appear ever to have considered the question of determining the minimum number of operations required to solve a given problem under a specified model of computation. For example, if we define (1) drawing a line and (2) drawing a circle as the primitive operations allowed under the straightedge and compass model of computation, **Algorithm Euclid** takes nine steps, whereas **Algorithm CO** takes only four steps. Its nonrobustness not withstanding, is **Algorithm CO** *optimal?* In other words, is four a lower bound on this problem? Is **Algorithm Euclid** the optimal *robust* algorithm? It is not difficult to argue that at least three steps are required. I conjecture that, in fact, four are always necessary.

This research suggests that perhaps the chaotic situation described here with respect to Euclid's second proposition exists also for his other propositions involving cases, and indeed for all of Greek mathematics. It may suggest a new way of examining old constructive mathematics and a new approach for historians of mathematics and philologists.

There are possible implications for education. It has been argued that Euclidean construction problems provide an excellent method of teaching high school students constructive proofs of existence theorems [3]. The work presented here suggests that Euclidean constructive geometry can be used as a medium for teaching modern concepts concerning the design and analysis of algorithms to high school students. For easy problems, the students can prove that Euclid's constructions are valid for all possible inputs. For more difficult problems, they can search for constructions that require fewer steps. Finally, for really challenging problems, they can search for constructions that require the fewest number of steps.

Acknowledgments

The author is grateful to Hossam ElGindy for translating Ibn al-Haytham's 11th-century Arabic manuscript [18]; to Mariza Komioti for translating Heiberg's definitive Greek edition [16]; to Kelly Lyons for bringing to his attention a copy of Peyrard's book at Queens University [28]; to Diane Souvaine, Sue Whitesides, and Chee Yap for discussions on this topic; and to Arnon Avron for providing some useful references.

References

1. A. V. Aho, J. E. Hopcroft, and J. D. Ullman, *The Design and Analysis of Computer Algorithms*, Reading, MA: Addison-Wesley, 1974.
2. A. Avron, Theorems on strong constructibility with a compass alone, *J. Geometry* 30 (1987), 28–35.
3. A. Avron, Construction problems—why and how? *Int. J. Educ. Sci. Technol.* 20 (No. 2) (1989), 273–287.
4. A. Avron, On strict constructibility with a compass alone, *J. Geometry*, 38 (1990), 12–15.
5. I. Barrow, *Euclide's Elements*, London: E. Redmayne, 1705.
6. P. Beckman, *A History of Pi*, New York: The Golem Press, 1971.
7. H. L. L. Busard, *The First Latin Translation of Euclid's Elements Commonly Ascribed to Adelard of Bath*, Pontifical Institute of Medieval Studies, 1983.
8. H. L. L. Busard, *The Latin Translation of the Arabic Version of Euclid's* Elements *Commonly Ascribed to Gerard de Cremona*, Leiden: E. J. Brill, 1984.
9. H. L. L. Busard, *The Medieval Latin Translation of Euclid's* Elements *Made Directly From the Greek*, Stuttgart: Franz Steiner Verlag Wiesbaden GMBH, 1987.
10. C. Clavio, *Euclidis Elementorum*, Francofurti: Jonae Rosae, 1654.
11. R. Courant and H. Robbins, *What is Mathematics?* Oxford: Oxford University Press, 1981.
12. W. Dunham, *Journey Through Genius: The Great Theorems of Mathematics*, New York: Wiley, 1990.
13. E. J. Dijksterhuis, *The First Book of Euclid's Elementa*, Leiden: E. J. Brill, 1955.
14. H. S. Hall and F. H. Stevens, *A Text-Book of Euclid's Elements*, London: Macmillan and Co., 1887.
15. Sir T. L. Heath, *The Thirteen Books of Euclid's Elements*, Cambridge: Cambridge University Press, 1928.
16. I. L. Heiberg, *Euclidis Elementa*, Lipsiae: B. G. Teubneri, 1883.
17. R. Honsberger, *Ingenuity in Mathematics*, New York: Random House, 1970.
18. Ibn al-Haytham, *On the Resolution of Doubts in Euclid's Elements and Interpretation of Its Special Meanings*, University of Istanbul, 1041 A.D., facsimile edition published by the Institute for the History of Arabic-Islamic Science at the Johann Wolfgang Goethe University, Frankfurt am Main, 1985.
19. G. J. Kayas, *Euclide: Les Elements*, Paris: Editions du Centre National de la Recherche Scientifique, 1978.
20. F. Klein, *Elementary Mathematics from an Advanced Standpoint: Geometry*, New York: Dover, 1939.
21. A. Kostovskii, *Geometrical Constructions with Compasses Only*, Moscow: Mir Publishers, 1986.
22. D. Lardner, *The First Six Books of the Elements of Euclid*, London: H. G. Bohn, 1861.
23. E. Lemoine, *Géométrographie*, Paris: C. Naud, 1902.
24. L. Mascheroni, *The Geometry of Compasses*, Pavia: University of Pavia, 1797.
25. G. Mohr, *The Danish Euclid*, Amsterdam, 1672.
26. G. R. Morrow, *Proclus: A Commentary on the First Book of Euclid's Elements*, Princeton, NJ: Princeton University Press, 1970.
27. D. Pedoe, *Geometry and the Liberal Arts*, New York: Penguin Books, 1976.
28. F. Peyrard, *Les Oeuvres D'Euclide, Traduites Littéralement*, Paris: C. F. Patris, 1819.
29. J. Playfair, *Elements of Geometry; Containing the First Six Books of Euclid*, Edinburgh: Bell & Bradfute, 1831.
30. F. P. Preparata and M. I. Shamos, *Computational Geometry: An Introduction*, New York: Springer-Verlag, 1985.
31. B. Russell, *The Autobiography of Bertrand Russell*, Boston: Little, Brown & Co., 1951.
32. M. E. Stark and R. C. Archibald, Jacob Steiner's geometrical constructions with a ruler given a fixed circle with its center (translated from the German 1833 edition), *Scripta Mathematica*, 14 (1948), 189–264.
33. J. H. Smith, *Elements of Geometry*, London: Rivingtons, 1879.
34. A. S. Smogorzhevskii, *The Ruler in Geometrical Constructions*, New York: Blaisdell, 1961.
35. H. M. Taylor, *Euclid's Elements of Geometry*, Cambridge: Cambridge University Press, 1895.
36. I. Todhunter, *The Elements of Euclid*, Toronto: Copp, Clark & Co., 1876.
37. G. T. Toussaint, Computational geometric thinking as kinesthetic thinking, presented at Conference on Thinking, Harvard University, August 1984.
38. R. Williams, *Elements of Euclid Explained and Demonstration in a New and Most Easy Method*, London: Globemaker, 1703.

School of Computer Science
McGill University
Montreal, Quebec H3A 2A7
Canada