To many students of beginning algebra, it must appear that algebra is just a matter of formal manipulations. But this is not true. As Middlemiss (1953) points out in the Preface, solving an equation $f(x) = g(x)$ to get $x = a$ is equivalent to proving

$$if \ f(x) = g(x), \ then \ x = a.$$ 

Furthermore, checking that $f(a) = g(a)$ is equivalent to proving the converse.

For the benefit of students just beginning to study algebra, this can be done without symbols:

* The work on which this document is based was supported in part by grant EQ41840 from the program Fonds pour la Formation de Chercheurs et l’aide à la Recherche (F.C.A.R.) of the Québec Ministry of Education and in part by grant OGP0023391 from the Natural Sciences and Engineering Research Council of Canada.
PROBLEM. Five more than three times a number is twenty; find the number.

Solution. Suppose that there is a number such that three times the number plus five is twenty.

Then, subtracting five from both “sides”,

three times the number is fifteen.

Hence, the number is five.

This proves: if five more than three times a number is twenty, then the number is five.

Checking, we see that five more than three times five is, indeed, twenty.

Doing some algebraic problems this way may be helpful to students who have trouble with the algebraic symbolism; for a discussion of some of the problems involved in teaching algebra see Herscovics (1989).

All this shows that algebra is not really about the use of symbols at all. Algebra really differs from arithmetic in two important ways:

I. In algebra, we use reasoning to solve problems indirectly when we cannot solve them directly.

II. We are concerned with general properties of numbers as well as particular problems (and we use reasoning to prove these properties).

However, the reasoning used in mathematics is not the same as in other fields. Reasoning in mathematics is more like the reasoning used among the sophists, as discussed by DeLong (1970) pp. 9-10:
... We do know that there existed a class of teachers who came to be known as sophists. These sophists would travel, much like wandering minstrels, and for a fee would teach their students how to speak persuasively on many different kinds of topics. Sophists were also prepared to defeat any opponent in a public argument. The competitiveness of such a spectacle must have been very keen and the arguments often dramatic, so that we can understand why the arrival of an important sophist in town was the occasion of much excitement and why sophists were often able to command large fees.

Protagoras, often considered to be the greatest of the sophists, would no doubt be thought a great thinker if his works had survived. He is best known for his saying that “man is the measure of all things” and his humanism probably exhibited itself in ways we consider uniquely modern. The following ancient story about him, although probably apocryphal, indicates the kind of verbal pyrotechnics of which the sophists were capable. Protagoras had contracted to teach Euathlus rhetoric so that he could become a lawyer. Euathlus initially paid only half of the large fee, and they agreed that the second installment should be paid after Euathlus had won his first case in court. Euathlus, however, delayed going into practice for quite some time. Protagoras, worrying about his reputation as well as wanting the money, decided to sue. In court Protagoras argued to the jury:

Euathlus maintains he should not pay me but this is absurd. For suppose he wins this case. Since this is his maiden appearance in court he then ought to pay me because he won his first case. On the other hand, suppose he loses the case. Then he ought to pay me by the judgment of the court. Since he must either win or lose the case he must pay me.
Euathlus had been a good student and was able to answer Protagoras’ argument with a similar one of his own:

Protagoras maintains that I should pay him but it is this which is absurd. For suppose he wins this case. Since I will not have won my first case I do not need to pay him according to our agreement. On the other hand, suppose he loses the case. Then I do not have to pay him by judgment of the court. Since he must either win or lose the case I do not have to pay him.

At this point, DeLong has the following footnote:

I have altered this story in inessential ways in order to bring out its logical form. For those who are interested in looking up the original story, see [Rolfe (1927) pp.] 404ff.

Imagine this case in a modern courtroom!

On the other hand, we use arguments like this in mathematics all the time. Thus, even if this story is not historically valid, it is a good illustration of the differences between reasoning in mathematics and reasoning in other areas.

Another difference between reasoning in mathematics and reasoning in other fields is that in mathematics we are interested in proofs of results which are intuitively obvious. This interest also goes back to the ancient Greeks, as Erik Stenius (1978) points on pp. 258-259:

This means that if “giving a proof” just means “proving what is not obvious from what is obvious” we may doubt that the Greek mathematicians were the first to prove theorems; but what seems very certain is that the Greek [sic] were the first to ask for proofs of “obvious” facts. So I state: The certainly original and revolutionizing idea of the Greek geometers was the endeavor to find proofs of “obvious” facts.
The difference between proving the nonobvious by means of the obvious and proving the obvious can be illustrated by the proof of the triangle theorem. The Pythagoreans used the well-known diagram (Fig. 1), which is found in many textbooks today. Now, when we have drawn the line $DCE$ parallel to $AB$ (taking “being parallel” in the etymological sense of “lying side by side” rather than in the Euclidean sense of “not intersecting”) the theorem becomes obvious from the figure, if it is only “seen” in the right way. For obviously the alternate angles $A$ and $ACD$ are equal, and likewise the alternate angles $B$ and $BCE$. If we then see $DCE$, on the one hand, as the sum of the three angles $ACD$, $ACB$ and $BCE$, and on the other hand as the sum of two right angles, then the theorem becomes “obvious”.

![Fig. 1](image)

But if we try to find a proof for the fact that the alternate angles mentioned are equal, then this special instance of proving the obvious happens to meet with difficulties which, as any mathematician knows, 2000 years later were to overthrow the whole conception of Euclidean geometry as stating “mathematical facts”.

Another example of this kind of proof (of the non-obvious from the obvious, which involves “seeing” a diagram the right way) comes from Plato’s *Meno*, where Socrates uses the following diagram to show that the area of a square whose
side is the diagonal of another square has twice the area of the other square.

Yet another example is the following proof of the Pythagorean theorem:
\[
\begin{align*}
&\text{Diagram 1:} \\
&\text{Diagram 2:}
\end{align*}
\]
I have heard this proof referred to as the “retiling proof” of the theorem.

According to Toretti (1978), p. 3, shortly after the time of Thales (c. 639—546 B. C.), the Greeks developed a new kind of proof in which the understanding did not come from a diagram but from an understanding of the meaning of the terms used. These were the first deductive proofs in mathematics. It is generally agreed that the oldest existing piece of deductive mathematics is preserved in Propositions 21—34 of Book IX of Euclid’s *Elements*, which are based on the definitions (especially Definitions 6 and 7) of Book VII. A diagram is given for each of these proofs (see, e.g., Heath (1956)), but that diagram is not the basis for understanding the proof. To understand the proof it is necessary to understand the meaning of the words used and to reflect on this meaning. As Toretti (1978) points out (p. 3):

Had they not adopted this method of exact, forceful, yet unintuitive thinking, Greek mathematicians could never have found out that there are incommensurable magnitudes, such as, for example, pairs of linear segments which cannot both be integral multiples of the same unit segment, no matter how small you choose this [i.e., unit] to be.

Toretti then goes on to quote B. L. van der Waerden (1961), p. 144:

When we deal with line segments which one sees and which one measures empirically, it has no sense to ask whether they have or not a common measure; a hair’s breadth will fit an integral number of times into every line that is drawn. The question of commensurability makes sense only for line segments which are objects of thought.
(This is an important point about mathematics that needs to be explained to beginning students.)

Yet the Greeks did find out that there are incommensurable magnitudes. This discovery is generally believed to have been made by Pythagoras or his followers. As Heath (1981) p. 91 puts it:

The actual method by which the Pythagoreans proved the fact that \( \sqrt{2} \) is incommensurable with 1 was doubtless that indicated by Aristotle, a *reductio ad absurdum* showing that, if the diagonal of a square is commensurable with its side, it will follow that the same number is both odd and even. [Here is a footnote referring to Aristotle, *Prior Analytics* i.23, 41 a 26l-27.] This is evidently the proof interpolated in the texts of Euclid as X. 117, which is in substance as follows:

Suppose \( AC \), the diagonal of a square, to be commensurable with \( AB \), its side; let \( \alpha : \beta \) be their ratio expressed in the smallest possible numbers.

Then \( \alpha > \beta \), and therefore \( \alpha \) is necessarily > 1.

Now \( AC^2:AB^2 = \alpha^2:\beta^2 \);

and, since \( AC^2 = 2AB^2 \), \( \alpha^2 = 2\beta^2 \).

Hence \( \alpha^2 \), and therefore \( \alpha \) is even.

Since \( \alpha : \beta \) is in its lowest terms, it follows that \( \beta \) must be *odd*.

Let \( \alpha = 2\gamma \); therefore \( 4\gamma^2 = 2\beta^2 \), or \( 2\gamma^2 = \beta^2 \), so that \( \beta^2 \), and therefore \( \beta \), is *even*.

But \( \beta \) was also *odd*: which is impossible.

Therefore the diagonal \( AC \) cannot be commensurable with the side \( AB \).

Note that although this proof could have been presented with a diagram of a square and its diagonal (as it is in Heath (1956), vol. III, p. 2), the diagram would add nothing to our
understanding. The principal way it is distinguished from the propositions in Book IX of Euclid is that it is an *indirect proof*.

How can we explain this transformation in Greek mathematics from a visual, intuitive approach to an abstract approach based on understanding the meanings of terms and reasoning about them? Our explanation should, if possible, give some indication of why it was the Greeks rather than some other people who carried out this transformation. It would also be useful if it gives us material that we can use to help explain to students what mathematics is all about.

We might begin by noting that the Greeks loved arguments, as DeLong (1970) points out in the story about Protagoras and Euathlus mentioned above. Furthermore, unlike most of the earlier ancient societies, they had no all-powerful class of priests. Thus, there were no positions that they were not prepared to question.

But this love of argument for its own sake is probably not enough to explain this change in Greek mathematics. For the change is more than just a change in the way truths are discovered. It also represents a change from a *practical* activity to a purely reflective one. The mathematics known from Egyptian and Babylonian sources was all closely associated with such practical activities as land surveying and architecture. It is only with the Greeks that it became a theoretical study without practical value, about which we can have a story such as that about Euclid given in Heath (1981), vol. I, p. 25:

The other story is that of a pupil who began to learn geometry with Euclid and asked, when he had learnt one proposition, ‘What advantage shall I get by learning these things?’ And Euclid called the slave and said, ‘Give him sixpence, since he must needs gain by what he learns.’
Now since all but fragments of pre-Euclidean Greek mathematics have been lost, we cannot be completely certain about the reason for this change in Greek mathematics. But there are a number of theories.

One of the theories is due to Arpád Szabó, of the Hungarian Academy of Sciences. In his (1978), he puts forward the view that this change is the result of influences from outside mathematics, and, in particular, from the Eleatic school of philosophy. This was the school of Parmenides (early 5th century B.C.) and his student Zeno (c. 490—c. 430 B.C.). The latter is known for his paradoxes of motion.

Szabó gives a description of the Eleatic philosophy in his (1978), pp. 217—218:

Their philosophy is distinguished by its rejection of practical empirical knowledge and of sense perception in general. Parmenides emphasizes that truth cannot be grasped by means of sense perception, which is misleading, but only by [sic.] reason ($\lambda\dot{o}\gamma\omega$). To get a clear idea of what he means by ‘reason’, let us take a look at one of his arguments; it asserts that what is ($\tau\omicron\delta\omicron\nu$) cannot have come into being and runs as follows: Suppose that what is did come into being, then it could only have come from what is or from what is not; there is no third possibility. Now if it had come from what is, it would already have been existent before it came into being; hence to say that it came into being in this way would make no sense. If, on the other hand, the claim is made that what is came from what is not, this leads immediately to a contradiction. What is can never have been the opposite of itself, what is not, and hence could not have come into being in this way either.
It is apparent that indirect arguments played a very important part in Eleatic philosophy. Without them it would not have been possible to establish such central doctrines as that there is no motion, no change, no becoming, no perishing, no space and no time. Of course, these doctrines contradict the evidence of our senses and are incompatible with empiricism, nonetheless the Eleatics, bolstered by their belief that reason was the only guide to truth, accepted them. Furthermore, the whole of Eleatic dialectic is nothing but an ingenious application of the method of indirect proof, which is why Aristotle considered Zeno to be the inventor of dialectic. … However, there is no real difference between Zeno’s dialectic and the arguments of Parmenides. The most noteworthy feature of both is their use of indirect proof.

As I remarked above, I believe that the influence of Eleatic philosophy was responsible for the rejection of empiricism and visual evidence in Greek mathematics, as well as for the introduction of indirect proof.

Szabó’s theory is that mathematics became a theoretical study as a result of an attempt to make the subject acceptable to the Eleatic philosophers. This involved difficulties, since these philosophers rejected multiplicity (which is clearly necessary for arithmetic) and space (which is necessary for geometry). According to Szabó, arithmetic was made acceptable to these philosophers by defining the unit in essentially the same way as Parmenides had defined what is and then presenting the other numbers as a new kind of multiplicity that does not occur in the world of the senses. They then tried formulating geometry as a purely theoretical, deductive science, but they never quite succeeded in making it acceptable to the Eleatics. This, according to Szabó, is why geometry eventually became a separate science.
Unfortunately, as Toretti (1978) says (p. 375, note 5), Szabó’s arguments are “not altogether convincing.” For example, Szabó argues that the theory of incommensurability (i.e., of irrational magnitudes) presented in Book X of Euclid’s *Elements* must have been completed before the time of Plato because in his dialogues Plato uses the technical terms involved in a way which indicates that they were well known to his audience. But Plato lived past his 80th birthday, from 428 B.C. to 347 B.C., and that is long enough for these technical terms to have become known to his audience during the course of his lifetime, or even of a part of it. H. B. Curry, whose student I was, had lived past his 62nd birthday when P. J. Cohen introduced the term “forcing” in his proof of the independence of the continuum hypothesis, yet long before the end of his life (at the age of 81) he was able to use this term before an audience of mathematical logicians and set theorists and assume that the audience would completely understand his meaning.

An alternative theory, which appears much more convincing, can be found in Knorr (1975). (This theory is really later than Szabó’s, since Szabó (1978) is a translation from a book published in German in 1969.) According to this theory, the transition to a theoretical discipline occurred as a part of the study of incommensurability. Knorr argues for this thesis by presenting a reconstruction of how this theory might have developed.

Knorr begins with the arithmetic of the later Pythagoreans in the fifth century B.C. He claims that the diagrams used in this study were rows of pebbles, so that the theorems on odd and even numbers in Book IX, Propositions 21-34 of Euclid’s *Elements* can be regarded as being based on seeing diagrams properly. For example, consider the following diagram for Proposition 21:
Compare this diagram with the text from Heath (1956):

PROPOSITION 21.

If as many even numbers as we please be added together, the whole is even.

For let as many even numbers as we please, \( AB, BC, CD, DE \) be added together; I say that the whole \( AE \) is even.

For, since each of the numbers \( AB, BC, CD, DE \) is even, it has a half part; \([\text{VII. Def. 6}]\) so that the whole \( AE \) also has a half part.

But an even number is that which is divisible into two equal parts; \([\text{id.}]\) therefore \( AE \) is even.

Q. E. D.

Note how much more Knorr’s diagram contributes to the understanding of this proof than the diagram given with the proposition in Heath (1956).

Another example is Proposition 22. Here are the two diagrams given by Knorr:
Now consider the text from Heath (1957):

**PROPOSITION 22.**

*If as many odd numbers as we please be added together, and their multitude be even, the whole will be even.*

For let as many odd numbers as we please, $AB$, $BC$, $CD$, $CE$, even in multitude, be added together; I say that the whole $AE$ is even.
For, since each of the numbers $AB, BC, CD, DE$ is odd, if an unit be subtracted from each, each of the remainders will be even; so that the sum of them will be even. But the multitude of the units is also even. Therefore the whole $AE$ is even. Q. E. D.

Note that an important part of understanding these proofs by seeing the diagrams the right way is understanding that the conclusions are valid not only for the specific diagrams provided, but for any diagrams which satisfy the hypotheses of the propositions.

To study the properties of square numbers, pebbles were arranged in squares. The results which can be proved this way include that the square of an even number is even, and is, in fact, a multiple of four. Consider this diagram:

```
  o o o  o o o
  o o o  o o o
  o o o  o o o
  o o o  o o o
```

Once it becomes that we can use essentially the same diagram for the square of any even number, the result is established. Similarly, the following diagram can be used to prove that the square of any odd number is odd and is also one more than a multiple of four:
In fact, once realize that no matter what odd number is used on each side, the four blocks in the corners are each a rectangle in which one side is one more than the other, we see that each of these blocks is even, so we can reach the stronger conclusion that the square of any odd number is one more than a multiple of eight.

These results on squares can easily be proved algebraically: $(2n)^2 = 4n^2$ and $(2n + 1)^2 = 4n(n + 1) + 1$. Note that since one of $n$ and $n + 1$ is even, their product is even.

If we apply these results to the problem of finding triples of integers which can serve as the legs and hypotenuse of right triangles in the light of the Pythagorean theorem, we come quickly to the following conclusions:

1. If the hypotenuse is even, so are both legs.

2. If the hypotenuse is odd, then one leg is even and the other is odd.

The key to these results is that it is impossible for the sum of two odd squares to be a multiple of four. Algebraically, we can see this as follows:

$$(8n + 1) + (8m + 1) = 4(2n + 2m) + 2.$$
Knorr suggests that the discovery of the incommensurability of the side and diagonal of a square occurred when these results were applied to the practical problem of finding the ratio of a diagonal of a square to its side, or, what amounts to the same thing, the ratio of the hypotenuse of an isosceles right triangle to a leg. This attempt leads to a puzzle. For if the hypotenuse is even, then so are both legs, and so we can bisect two sides and obtain a new right isosceles right triangle half the size of the original one. We can repeat this process each time we get an isosceles right triangle with an even hypotenuse. Hence, if we start with a given isosceles right triangle and repeat this process, we must eventually get an isosceles right triangle with an odd hypotenuse. For this triangle, one leg must be even and the other odd. But since the triangle is isosceles, the legs are equal, and so we must conclude that an even number is equal to an odd number, which is impossible. This puzzle was resolved when it was realized that the argument constituted an indirect proof that the diagonal and side of a square have no common measure. Note that there is no way to understand this indirect proof by “seeing” a diagram the right way. According to Knorr, it was at this point that indirect proofs and proofs that depend on understanding the meaning of the terms involved and reasoning about them were introduced into Greek mathematics. (According to Knorr, this happened about 430 B.C. This is too late for the introduction of indirect proofs to have come directly from the Eleatics, but they may have been taken from other philosophers and thus indirectly from the Eleatics.)

It is tempting to think that this discovery caused a major crisis in mathematics and philosophy, for the Pythagoreans had assumed that everything in the universe could be represented in terms of (whole) numbers. Furthermore, in mathematics itself, ratio was defined for numbers, so the general theory of similar figures was shown to be without a proper foundation. But this is not what happened. Mathematicians and philosophers continued as before; in particular,
geometers continued to use the theory of similar figures. (This is just what happened at the beginning of this century when paradoxes were discovered in the foundations of set theory; most mathematicians continued to use set theory as though nothing had happened.) It appears that the only lasting change in mathematics caused by this discovery was the introduction of proofs by contradiction and proofs that involved abstract reasoning rather than seeing diagrams.

According to Knorr, the next development occurred when Theodorus of Cyrene (active from 410 to 390 B.C.) decided to study incommensurability. Knorr proposes a reconstruction of how he might have been able to prove the incommensurability with a unit of the sides of squares of area 3, 5, 6, 7, 8, 10, 11, 12, 13, 14, and 15; this reconstruction does not use any more geometry or number theory than was involved in the proof that the diagonal of a square is incommensurable with its side. The method of proof used by Knorr fails for a square of area 17 (because 17 is, like any square of an odd number, one more than a multiple of 8).

This reconstruction makes it plausible that Plato’s dialogue *Theaetetus* can be interpreted as representing the history of the development of this mathematical theory. In the dialogue, which is set in 399 B.C., Theodorus’ pupil Theaetetus (414—369 B.C.) recounts a lecture by Theodorus presenting the above theory, and stating that Theodorus stopped at 17. (Knorr maintains that the text should be read to mean that Theodorus ran into trouble at 17.) The dialogue goes on to have Theaetetus assert that (in effect) the square root of every integer which is not a perfect square is irrational. Knorr reconstructs how Theaetetus (along with Archytas of Tarentum, active after 390 B.C.) managed to prove this result; the proof required developments in both number theory and geometry. Furthermore, since the number theory was being applied to geometry, numbers began to be represented by line segments instead of the pebble diagrams. (Knorr says that this began in the time of Theodorus.) This effectively disguised the fact that the proofs had originally been “seen”
to be correct by looking at diagrams, but by this time mathematics was already largely a theoretical discipline.

One of the developments attributed to Theaetetus was a definition of ratio and proportion for incommensurable magnitudes. (According to Knorr, Theodorus did not have such a definition and did not need it. Knorr refers to the the theory as a theory of proportion, but Fowler (1979) says that it was really a theory of ratio. The word "ratio" is not defined in Euclid.) This is not the theory of Book V of Euclid’s *Elements*, but an earlier theory of which we no longer have any contemporary manuscripts, and hence it has had to be reconstructed by historical detective work.

Theaetetus worked at Plato’s Academy in Athens. After he died, Eudoxus (395—340 B.C.) joined the Academy and the mathematicians working there. By this time, according to Knorr, Plato was pushing the mathematicians to put their theories in what we would call an axiomatic form, in which all the theorems used in proofs (except, of course, for the basic axioms, postulates, and definitions) is rigorously proved. According to Knorr, Eudoxus decided to prove a result which had previously been taken for granted, namely that if $\frac{A}{C} = \frac{B}{C}$ then $A = B$. This proof turned out to be extremely difficult in terms of the definition of ratio introduced by Theaetetus. Thus, when Eudoxus found that he could use what we know as Definition 5 of Book V of Euclid’s *Elements* (which is a definition of proportion, not ratio) to give a simpler proof of the same result, he was able to replace Theaetetus’ theory of ratio with a simpler theory. Eventually, Theaetetus’ theory was all but forgotten.

This historical reconstruction by Knorr shows mathematics becoming a theoretical discipline involving an axiomatic approach as the result of the mathematical (not originally philosophical) demands of the main theory being studied at the time, namely incommensurability. Furthermore, the above summary of the reconstruction down to the indirect proof of the incommensurability of the side and diagonal
of a square can easily be presented to beginning algebra students.

This suggests some changes in a beginning algebra course. I propose beginning with the example without symbols from the beginning of this paper (or one like it) to make the point that algebra differs from arithmetic in using reasoning to solve problems indirectly and prove general results. I would then do elementary area formulas using diagrams. The formula for the area of a rectangle can be made plausible by counting squares; I would start with rectangles whose sides are whole numbers, and then do rectangles with sides involving rational fractions by changing the size of the unit. (I would leave out any mention of irrational sides at this stage.) The formula for the area of a parallelogram can be justified by the following diagram:

![Parallelogram Diagram]

Similarly, the following diagram will justify the formula for triangles:

![Triangle Diagram]

The Pythagorean Theorem can then be proved by the diagrams given earlier in this paper. It is then possible to use
the pebble diagrams to prove elementary facts about even and odd numbers, and to get to the facts about Pythagorean triples and Knorr’s version of the proof of the incommensurability of the diagonal and side of a square.

At this point, I would take time to criticize the earlier treatment of rectangles with rational sides. The object would be to develop an awareness that there are cases that are not covered by that treatment. Other subjects involve criticism of previous work in early courses, and I think mathematics courses would benefit from the same thing.

After this, I would proceed with the use of algebraic notation to solve simple equations and the rest of the traditional curriculum.

Later in the course, it might be worth taking up the question of the area of a circle. I would approach the result in the form Archimedes proved it, namely, that the area of a circle is equal to the area of a triangle whose base equals the circumference of the circle and whose altitude equals the radius. (See Heath (1897), p. 91ff.) I would introduce the following diagram:
This diagram shows a circle with one square inscribed and another circumscribed. Each square is made up of four triangles whose bases form the perimeter and whose third points are the center of the circle. If one looks at this diagram long enough, one will see that the area of each square is equal to the area of a triangle whose altitude is the altitude of any one of the four triangles and whose base is the perimeter. In the case of the inscribed square, the perimeter is less than the circumference and the altitude is less than the radius; in the case of the circumscribed circle, the perimeter is greater than the circumference and the altitude equals the radius. If we double the number of sides of the inscribed and circumscribed polygons, the inequalities mentioned above are preserved, as is the formula for the area of the polygon. Also, the perimeters and areas are clearly closer to each other. So if we think of continuing to double the number of sides, we must be narrowing in on the area of the circle. This makes Archimedes’ theorem extremely plausible. However, the problem of really nailing down a proof involves the concept of a limit, and so a discussion of this problem can help prepare the way for more advanced courses.

References

Herscovcics, Nicolas (1989). Cognitive obstacles encountered in the learning of algebra. In Sigrid Wagner and Car-


