1 Ingredients and Tools

1.1 Definitions and Undefinitions

Richness in literary expression depends in part upon the variety of meanings that words have, upon the vagueness and ambiguities of the ideas expressed. This cannot be allowed in mathematics, where clarity and precision of expression rather than richness is needed. In mathematics, words must be carefully defined in such a way that the definitions satisfy certain prescribed criteria: simplicity; noncircularity; unique characterization. This is not always achieved, but it should at least be an ideal at which to aim. Consider the following illustration.

Let us define a "pencil" as "a sharp tool used for writing." This satisfies the criterion of simplicity, which means that any definition should use ideas and words that are as simple or simpler than the idea or words being defined. To anyone who does not know what a pencil is, the following definition would be incomprehensible: "a cylindrical instrument tapering to a conical point—as employed by an amanuensis." A definition using words more complicated than the one being defined is not of much use.

But even in a definition using simple words there is an inherent problem closely connected with the criterion of noncircularity. It is simply not possible to define every word without getting trapped either in an infinite process of defining or in circularity. Dictionary definitions are basically circular. In the suggested definition of "pencil," if one does not know the meaning of "tool" one might look it up to find it is an "instrument" or "utensil." And since it is not likely that anyone looking up the word "tool" would know the meaning of "instrument" or "utensil" either, these words would have to be looked up also. Sooner or later, a word in this chain would be defined as a "tool."

To see how serious this problem might be, consider what would happen if you wanted to use a dictionary in a foreign language, knew nothing of the foreign language, and had available only a dictionary written totally in that language. An immigrant to our shores, for example, would find any standard English language dictionary useless to him if he knew absolutely no English.

To avoid this problem it is necessary to choose some words as primitive, or undefined words, words in terms of which all the other words of the system may be defined.

So, the definition of "pencil" as a "sharp tool used for writing" satisfies the criterion of simplicity but fails to comply with noncircularity. It also fails to satisfy the third criterion, and this proves most often troublesome in mathematics.

When it is said that a suggested definition does not uniquely characterize that which is being defined—in this case a pencil—this means that it is equally applicable to other things—in this case to pens, or for that matter to quills.

To sum up, any good definition must be expressed simply, must be noncircular and must uniquely characterize that which is being defined.

With this in mind let us look at the first three definitions in Euclid's *Elements:*

1. A *point* is that which has no part.
2. A *line* is length without breadth.
3. A *straight line* is that which lies evenly with the points on itself.

These are not good definitions. The first two do not characterize that which is being defined. The first definition might as well say that *courage* is that which has no part; it might say that a *ghost* is that which has no part (or, for that matter, is length without breadth). For centuries commentators, both critics and defenders of Euclid, have discussed and argued about his definitions. Proclus (fifth century A.D.), perhaps his best commentator and defender, says, "like the *now* in time and the *unit* in number, a point is that which has no part, but in the subject matter of geometry, a point is the only thing which has no part." He thus argues that within the context this definition uniquely characterizes what is being defined. It might be argued without end whether or not these two are good definitions, but in the case of the third definition the problem is rather cut-and-dried. It is circular and should be avoided. It is impossible to define "lies evenly" without using the concept of "straight" or something synonymous with it.

In addition to the foregoing, the first two definitions have been criticized for surreptitiously introducing philosophical and physical problems that do not belong to geometry as an abstract formulation. This criticism is consistent with the spirit of Euclid, who would frown upon his work as a "practical" science. In fact, some over-enthusiastic defenders of Euclid claim that Euclid did not mean statements (1) and (2) to be definitions but was merely pointing out the abstract, nonspatial characteristics of his work. If so, if he does not mean to define point and line, he is in this respect a modern.

In modern systems the problem of circularity, as well as the problem just mentioned, is avoided by taking many words as undefined. Among these, two types are distinguished.

The *technical* terms. These vary from subject matter to subject matter. In geometry, such words as "point," "line," "congruent," "between," might be considered as the primitive or undefined terms of the system being considered. It is possible that some other system of geometry might choose other undefined terms, but those that are more or less specific to geometry are the technical terms.

The *logical*, *language*, or *universal* terms. These are words such as "all," "every," "any," "there exists," "at least one," "at most one," "only," "the," "although," and so on. The list goes on indefinitely, but those mentioned are the words that occur most frequently in mathematics. When a system is devised for geometry, such words as "one" and "two" are usually taken as part of the universal language.

The technical and universal terms listed in the preceding paragraphs are those which remain *undefined*. There are also technical and universal terms which are *defined;* in mathematics, definitions are usually limited to technical terms. It may surprise many readers to hear that definitions are not really necessary. This is indeed so, but we would be very verbose and would be more likely to be inaccurate and even contradictory without them. If the very same words are going to be used to describe a concept every time it is expressed, it may as well be abbreviated; that is all a definition is. If, on the other hand, the same words are not going to be used to express a concept each time it occurs, the chance for introducing contradictions increases.

The reason that mathematics leans so heavily on symbolism as a means of expressing a concept is not merely that it is a convenient shorthand, but also because it allows the mathematician to avoid expressing the same concept in words. Ordinary language words are used very sloppily, and this is true even when a mathematician uses them. As an example, consider the word "circumference" used in reference to a circle; the commonly accepted definition today is that it is the *distance* around the circle, yet one often reads and hears the word used as if it meant the circle itself, as when we speak of "a line cutting the circumference." Another phrase which frequently occurs is "the area of a circle." Yet the commonly accepted definition of a circle is: the set (or locus) of points in a plane equidistant from a fixed point called the center. Obviously, a locus of points has no area. And if by a "circle" one were to mean that portion of the plane enclosed by the locus of points, it would then be false to say that a line may cross a circle in at most two points, or that two circles can cross in at most two points, for they might have an infinite number of points in common.

Our use of symbolism will be limited. We will introduce it when convenient and avoid it when we think it best. Instead, we shall introduce many, many definitions, attempting always to satisfy the three stated criteria and to use the words precisely as they are defined. In the case of undefined technical words, the axioms of the system will restrict how they will be used.

EXERCISES 1.1

1. Compare the use of "=" in the following four cases and explain any differences.

(a) $\cos^2 x + \sin^2 x = 1$
(b) $1/2 = 2/4$
(c) $2x + 5 = 11$
(d) angle A = angle B

2. In a beginning calculus class, the instructor has covered the blackboard with a "delta-epsilon" proof for the uniqueness of limits. When he finishes, a student asks him "If the lim $f(x)_{x \to a} = A$ and the lim $f(x)_{x \to a} = B$, and if things equal to the same thing are equal to each other, is it not possible to say simply that $A = B$? Why cover the board with all that stuff?"

Does the student have a point? Explain.

3. An example of a good definition is: A square is a four-sided figure bounded by straight lines; its opposite sides are parallel; and all four of its sides are equal. Answer true or false and explain your answer.

1.2 Axioms

What precisely is an axiom? One might think that this is a good place to start with a "simple, noncircular and uniquely characterizing" definition. If one is going to speak about axioms, why not immediately define what the word means? Unfortunately, it is not that easy.

Suppose we use the dictionary definition: an *axiom* is a self-evident or universally recognized truth, accepted without proof. The trouble with this is that modern mathematics would not accept this as a definition. The last part "accepted without proof" is all right, but the rest is not. To appreciate what has brought about this situation, let us look at the statements that Euclid chose as axioms.[1]

1. That a straight line may be drawn from any one point to any other point.
2. That any finite straight line may be produced to any length in a straight line.
3. That a circle may be constructed with any center, at any distance from that center.
4. That all right angles are equal to one another.
5. That, if a straight line falling on two straight lines makes the two interior angles on the same side of it taken together less than two right angles, the two straight lines, if produced indefinitely, meet on that side on which are the angles less than two right angles.

While Euclid *may* have regarded points and lines as mere abstractions, as ideals of reality that can only be approximated in our spatial world, it is unlikely that he regarded his "points" and "lines" as mere variables or undefined terms. His "points" and "lines" were in some way related to what is drawn on paper and the blackboard. His axioms were regarded as truths of the world we live in, not "mere" assumptions of someone's fancy.

As late as the eighteenth century, the great German philosopher, Immanuel Kant, built a tremendous philosophical structure around statements he called "apriori synthetic," statements, he said, such as those in geometry, which, while rooted in experience, are yet truths of an unquestionable nature.

The axioms of geometry were for two thousand years regarded as "self-evident universal truths." Is it any surprise, then, that it was not until the nineteenth century that even the possibility of a non-Euclidean geometry was put forth? Is it any surprise that Karl Friedrich Gauss (1777–1855), after developing and proving many of the theorems of a "new" geometry, kept his discoveries secret for over thirty years in fear that their revelation might harm his reputation as a mathematician? For

[1] All quotations from Euclid come from *The Thirteen Books of Euclid's Elements* translated with introduction and commentary by Sir Thomas L. Heath, 2d ed. (New York: Cambridge University Press, 1926. Reprinted by Dover Publications, Inc., 1956.)

if Euclid's geometry was the true geometry, any other would be either nonsensical or false.

Today, with the discovery and acceptance of various consistent non-Euclidean geometries, the status of axioms has undergone a radical change. To see just how drastic the change has been, contrast the views of Kant with those of Bertrand Russell (1872——). "Mathematics," Russell has said, "is that subject in which we do not know what we are talking about, or whether what we are saying is true." Even those who do not agree with such a view will recognize that it serves to emphasize an important aspect of the structure of mathematics; namely, that mathematics, instead of being the absolute science, is really the science of "if . . . then . . .". But that a statement such as Russell's can be made seriously needs explanation.

Although the way in which mathematicians regard statements has evolved over a number of years and has been influenced by many minds, Hilbert, because of the great respect he commands, has profoundly affected the course of mathematics in the twentieth century. Two of the concepts he incorporated into his axiom system for Euclidean geometry must be noted: some terms must be taken as undefined; axioms are mere assumptions about these terms.

From this beginning it is not a big step to believing that, at least where mathematics itself is concerned, the truth or falsity of the axioms is not a crucial question. For mathematics the crucial problem becomes one of "if . . . then . . ."; that is, *if* we assume such and such axioms, *then*, what follows from them. Furthermore, strictly speaking, if we regard the technical terms as undefined terms, as mere variables, then the axioms are themselves in a sense variable, or "open," sentences. As such, they cannot be said to be either true or false.

If one says, "He was the greatest mathematician who ever lived," one is not really making a *statement*—that is, to such a collection of words the labels "true" and "false" do not apply. If the sentence is changed to "Beethoven was the greatest mathematician who ever lived" or "Gauss was the greatest mathematician who ever lived," then the sentence becomes a *statement*, a collection of words to which the labels "true" or "false" may be applied. A similar comment can be made for such sentences as, "x is a whole number," "every dabba is a set of abbas." It is because of this quality of the "open" sentence that a statement such as Russell's can seriously be made.

When one begins to view axioms in this way, a question naturally arises. Of the infinite variety of such statements, how does one determine which should be used for axioms? There is no technique, no mechanical process, to help us. Assuming that one is not going to put together a meaningless, or random, collection of statements, assuming that there is some underlying unifying concept and some set of statements that one wishes to prove, the determination of the precise set to be chosen as axioms is a creative act. Yet this is only the beginning. One must then determine whether or not these axioms satisfy certain properties among themselves, and what the relationship between them and other statements, called "theorems," is. We will consider these and other questions in the next two chapters.

EXERCISES 1.2

State whether the following are true or false. Justify your answer.

1. In an axiomatic system, *every* word must be carefully defined.
2. In an axiomatic system, some *technical* words *must* be defined.
3. In an axiomatic system, if some words are defined, some must remain undefined.
4. In modern mathematics the trend is to regard axioms as self-evident truths.
5. If we regard an axiom as an open sentence, it is neither true nor false.

1.3 Logic

From a few definitions, and a few axioms, Euclid derived many "theorems"—statements which are said to "follow" from the others. Although Euclid's axioms and theorems may not be of the "if . . . then . . ." form, most of them can be put into that form. Furthermore, even if some of the statements cannot be so rephrased, the overall "if . . . then . . ." quality of the system remains. For he is saying that *if* such and such statements are granted, *then* such and such statements follow. This process of reaching conclusions from axioms is called *deduction*. The relationship that holds between the statements taken as axioms and those which are deduced from them is called *implication*. We say that the axioms *imply* the theorems.

When we consider *statements* of the form "if . . . then . . .", usually called conditional statements, we shall, contrary to common procedure, refrain from using "implies." In this context, its use tends to be confusing, because a conditional statement is just that, a statement. Given a statement P, and another statement Q, then we can define a new statement as follows:

P	Q	P → Q
T	T	T
T	F	F
F	T	?
F	F	?

(a)

P	Q	P → Q
T	T	T
T	F	F
F	T	T
F	F	T

(b)

It is customary to take Table (b) as the *definition* of "→", which can be read "arrow." Thus we might read "$P \rightarrow Q$" as "P arrow Q." More often, however, it is read as "If P then Q," which helps explain why Table (b) is chosen as its definition. Consider the following typical conditional statement: If L is a line, then there exists a point not on L. This statement says something about the existence of points under the hypothesis that L is a line; it says nothing about the existence (or nonexistence) of points if L is not a line. Looked at another way, we could say that such a statement is false under exactly one set of circumstances: L is a line, and there exists no point not on L. Ordinarily then, if L were not a line, we would say that the rest of the statement simply would not arise; that is, there is no restraint upon Q when P is false.

This is compatible with Table (a). The reason for choosing Table (b) is that logic depends on two classical laws: *The Law of Contradiction*—no statement can be both true and false and *The Law of Excluded Middle*—any statement must be either true or false.

As can be seen in Table (b), there are four possible truth combinations of P and Q. If "$P \rightarrow Q$" is to depend upon the truth of the component statements, it must have a value of either true or false in the last two lines, and because we are attempting to define a conditional statement which can be false only in the case of line two, we give the last two lines values of "T".

If we now read "$P \rightarrow Q$" as "P implies Q", which is often done in mathematics, and still interpret P implies Q to mean "Q follows from P," we can get all kinds of strange statements such as "'5 + 5 = 11' implies, 'The moon is made of spaghetti sauce.'" And who, other than a "brainwashed" student, would admit that "the moon is made of spaghetti sauce" *follows from* "5 + 5 = 11"?

Another reason for restricting our use of "implies" is that the relationship between axioms and theorems is just that, a *relation*, whereas "→" is not a relation but a type of *operation*. Thus, "→" forms a new statement, whereas "implies" does not; it "talks about" statements. Analogous to this is the difference between the "addition" operation and the "less than" relation. When two numbers are added, a new number is obtained, but when we say "2 is less than 3" we do not obtain a new number; we are expressing a relation, we are "talking about" numbers.

There are good reasons why these two concepts have tended to fuse. One is that both "$\rightarrow$" and "implies" are read in English as "if . . . then . . .". Closely connected with this is the fact that both have similar relations with the truth values of the statements involved. Just as "$\rightarrow$" is false when P is true and Q is false, so, if P implies Q, then one combination of truth values is ruled out; it is impossible for a true statement to imply a false statement. We wish to avoid saying, however, that *any* false statement *implies* any statement.

To sum up: we shall consider *conditional statements* as defined by Table (b). We shall consider the *relation implies* to be an undefined relation of our universal language, and shall resort to citing many, many illustrations to indicate what we mean when we say that one statement "follows" from another.

Two other operations must be introduced in this section: "not" and "and."

If "P" is a statement, "not-P" is its denial as defined by Table (c).

If "P" and "Q" are statements, then "P and Q" is the simple conjunction as defined by Table (d).

P	not-P
T	F
F	T

(c)

P	Q	P and Q
T	T	T
T	F	F
F	T	F
F	F	F

(d)

P	Q	not-Q	P and not-Q
T	T	F	F
T	F	T	T
F	T	F	F
F	F	T	F

(e)

The definitions of "$\rightarrow$", "not" and "and" can now be used to construct new statements and their denials. For example, "P and not-Q" is the denial of "$P \rightarrow Q$", as shown in Table (e).

It can be seen that Table (e) has F and T exactly interchanged with Table (b), and hence is its denial as defined by Table (c).

This use of "denial"—that given two statements, if one is true the other is false and if one is false the other is true, is called *contradiction.* The two statements are said to be *contradictory.* A statement such as "P and not-P" is said to be a *self-contradiction.* Whenever we want the contradiction of a statement we shall ask for *the* denial.

There is another combination that is *a* denial but not the contradiction. To illustrate, consider the statement "Today is Tuesday." A denial of this might be "Today is Wednesday." These statements deny each other but do not contradict each other, for they may both be false. Such denials are called *contraries;* they are statements that cannot both be true but that may both be false.

Associated with the conditional statement are two others worthy of specific mention: Given "If P then Q"

1. "If Q then P" is called the *converse.*
2. "If not-Q then not-P" is called the *contrapositive.*

We shall also make statements of the form "P if and only if Q," abbreviated "P iff Q," which is to be the same as the conjunction of a conditional statement and its converse. *All definitions, whether explicitly given as such or not, may be regarded as "iff" statements.*

Rather than attempt to condense a short course in logic into a brief section, we have discussed only those topics that will be required later and whose discussion within the text would have been too much of a digression. A few others will be discussed as they arise: partial converses; universal statements and their denials; and proofs, indirect and direct.

EXERCISES 1.3

1. Show that $Q \rightarrow P$ has the same truth table as not-$P \rightarrow$ not-Q.
2. Construct truth tables for: (a) $P \rightarrow Q$ (b) $P \rightarrow$ not-Q (c) P and Q (d) P and not-Q and compare.

Are $P \rightarrow Q$ and $P \rightarrow$ not-Q contradictory? Are any of the statements above contradictory to each other?

3. Write the converse of not-$Q \rightarrow$ not-P.
4. Write the contrapositive of not-$P \rightarrow Q$.
5. Write the converse and contrapositive of each of the following.

(a) If two lines are parallel, they have no point in common.
(b) If two angles are congruent, they are right angles.
(c) If a triangle has three equal angles, it is equiangular.

6. Is the converse of every definition, expressed in "if . . . then . . ." form, true? Explain.

7. Is the converse of every true conditional statement a true statement? Explain.

8. Is the converse of every false conditional statement a false conditional statement? Explain.

9. (a) Does "P only if Q" mean "$P \rightarrow Q$" or "$Q \rightarrow P$"?
   (b) Does "P if Q" mean "$P \rightarrow Q$" or "$Q \rightarrow P$"?

10. Write in "if . . . then . . ." form:

(a) Two lines are parallel only if they have no point in common.
(b) Two lines are parallel if they have no point in common.

11. Separate into two statements of the "if . . . then . . ." form:

(a) Two sides of a triangle are congruent iff the angles opposite them are congruent.
(b) Two lines are perpendicular to a third line iff they are parallel.
(c) Two lines, intersected by a third, are parallel iff the alternate interior angles are congruent.

1.4 Sets

The concept of a *set* shall also remain undefined. When we say that a set is any collection of objects we are not defining it. We could just as easily use the words *collection*, *class*, or *group*. But the word *set* is customary in mathematics, while the word *class* is used in philosophy, and no mathematician would use the word *group*, which has another special meaning. The modern mathematical theory of sets is usually credited to the German mathematician, Georg Cantor, who, in attempting to define the word "set," became involved with an intricate problem. It is because of this that mathematicians no longer attempt to define the word.

There is a standard notational shorthand that goes with the theory of sets. We shall for the most part avoid its use in this book but, because it is a common and convenient timesaving device in classroom lectures, we shall introduce the basic concepts and symbols of set theory.

If a set, denoted by S, is a collection of things, say a, b, and c, it is written:

$$S = \{a, b, c\}$$

Two sets are said to be *equal* iff they have the same elements. Thus, the set $A = \{a\}$ and the set $B = \{a, a, a, a\}$ are to be regarded as the same set. Two sets which are not equal are said to be *distinct*.

To say that a *belongs to S*, or *is an element* of S, one writes:

$$a \in S$$

To deny this, one writes:

$$a \notin S$$

If A is any set such that all of its elements are also elements of S, then one says that A is *a subset of S*. From the example S given before, if $A = \{a, b\}$ then A is a subset of S, which is written:

$$A \subseteq S$$

By definition, it follows that $A \subseteq A$, and $S \subseteq S$, or, any set has itself as a subset. If one wishes to write that A is a subset of S but not equal to S, this can be written:

$$A \subset S$$

and read "A is a *proper subset* of S." Many authors use this symbol for subset and have no symbol for proper subset.

If one wishes to talk about all the elements of a set, one might merely list them. But suppose that this cannot be done. Then the following notation is useful:

$$S = \{x | x \text{ is a point}\}$$

which is read "the set of all x such that x is a point" or more naturally "the set of all points."

If A and B are sets, then the elements common to both sets, that is, the set of elements which belong to both A and B, is called the *intersection*. To denote this, we use the notation:

$$A \cap B$$

It may happen that two sets have no elements in common; in such a case the sets are called *disjoint*, symbolized by:

$$A \cap B = \phi$$

where ϕ is called the *null*, or *empty* set, *the* set that has no elements. From the definition of subset, it follows that the null set is a subset of any set.

Because in this book we are going to be considering lines as sets of points, we shall adopt the following convention. Whenever we use the verb "intersects" we shall mean a nonempty intersection. So we shall never say that two sets intersect when in fact they are disjoint.

If A and B are sets, then the set of all elements that belong either to A, or to B, or to both, is called the *union*. This is denoted by:

$$A \cup B$$

For example, if $A = \{a, b, c\}$ and $B = \{a, b, d\}$ then $A \cup B = \{a, b, c, d\}$. Observe that there is no need to write $A \cup B = \{a, a, b, b, c, d\}$.

Further comments on this topic will be incorporated into the text.

EXERCISES 1.4

1. If $S = \{1, 2, 3, 4, 5, 6\}$ $L = \{1, 3, 5\}$ $M = \{2, 4, 6\}$ $N = \{4, 6\}$ $P = \{6\}$, find:

(a) $L \cup M$
(b) $L \cap M$
(c) $N \cap N$
(d) $M \cup M$
(e) $L \cap P$
(f) List all the subsets of N; of M.

2. Using L, M, N of exercise 1, which of the following hold?

(a) $L \cup (M \cap N) = (L \cup M) \cap (L \cup N)$
(b) $L \cap (M \cup N) = (L \cap M) \cup (L \cap N)$
(c) $L \cup (M \cap N) = (L \cup M) \cap N$
(d) $L \cap (M \cup N) = (L \cap M) \cup N$

3. If $X \cup Y = X$ and $X \cap Y = X$, then $X = Y$. True or false. Explain.

4. For every set A, $A \cup \phi = A \cup A = A \cap A$. True or false. Explain.

5. For every set A, B, $(A \cap B) \subset A$ and $(A \cap B) \subset B$. True or false. Explain.