The Thrills of Abstraction

P. R. Halmos

My wife and I were invited to a party recently, a party attended by four other couples, making a total of ten people. Some of those ten knew some of the others, and some did not, and some were polite, and some were not. As a result a certain amount of handshaking took place in an unpredictable way, subject only to two obvious conditions: no one shook his or her own hand and no husband shook his wife's hand. When it was all over, I became curious and I went around the party asking each person: "How many hands did you shake? . . . And you? . . . And you?" What answers could I have received? Conceivably some people could have said "None", and others could have given me any number between 1 and 8 inclusive. That's right, isn't it? Since self-handshakes and spouse-handshakes were ruled out, 8 is the maximum number of hands that any one of the party of 10 could have shaken.

I asked nine people (everybody, including my own wife), and each answer could have been any one of the nine numbers 0 to 8 inclusive. I was interested to note, and I hereby report, that the nine different people gave me nine different answers; someone said 0, someone said 1, and so on, and, finally, someone said 8. When it was all over, my curiosity was satisfied: I knew all the answers. Next morning I told the story to my colleagues at the office, exactly as I told it now, and I challenged them, on the basis of the information just given, to tell me how many hands my wife shook.

I mention this problem not because it is what I want mainly to say (but I will parenthetically whisper that it is a legitimate, bona fide problem, and my colleagues could have solved it), but because it illustrates, sort of, one of the basic mathematical principles that I do want to discuss. Here is another question, that leads to another basic mathematical principle: is there a number with the property that when it is multiplied by itself five times the result is the same as if we had just added 2 to it? Anyone who did not manage to remain completely innocent of high-school algebra will recognize the question as a non-symbolic way of asking whether the equation $x^5 = x + 2$ has any solutions.

Numbers, Phonemes, and Species

Here is one more question: what is there in common between the biologist's concept of a species, the linguist's concept of a phoneme, and the mathematician's concept of a number?

I'll turn to the questions in reverse order, but before doing that I'd like to describe the leitmotif of the whole discussion.

What is a black hole? I don't know, but I get a vague idea every now and then from a casual article in a newspaper or a magazine. It seems to be something very very heavy—so heavy that nothing that once enters its domain of gravitational attraction can ever escape from it, not even light, and, as a result, it is something that we can never perceive with any of our senses—something we can never see, hear, smell, taste, or feel. It has a measurable influence on some of the world we can perceive, but it itself is an abstraction. I probably said that all wrong, but I feel that even my vague and erroneous notion of a black hole is worth knowing—when I learned it, my soul grew, I became richer. I saw a vista I had never dreamt of before, my imagination was stimulated in a new way.

All parts of human intellectual endeavor have their abstractions: the economist's utility, the psychologist's id, the chemist's molecule, the biologist's species, the linguist's phoneme, and, of course, the mathematician's number—all these are abstractions, and each is a seminal part of the field it belongs to. When I asked, however, what species, phoneme, and number have in common, I wasn't just leading up to the shallow answer that they are all abstractions. There is more to the question than that.

A dictionary might define a phoneme of a language as "a smallest unit of speech that distinguishes one word from another". That's too quick, too shallow, too simplistic, but it's a beginning of a definition. An example will help to clarify the issue. If I replace *b* by *m* in "bat", I get another English word, "mat", that means something completely different; that's why *b* and *m* belong to two different English phonemes.

Consider, on the other hand, the words "stone" and "tone". Does everybody realize that the *t* sounds different in those words? In "tone" it is aspirated, and in "stone" it is not—by which the linguist means that if I hold a slip of paper two or three inches from my lips and then say "tone", the paper will move, but if I say "stone", it will not. There are languages (I believe Hindi is one of them) in which the replacement of an unaspirated *t* by an aspirated one can change meaning (just as the replacement of *b* by *m* changes the meaning of "bat"). In English, however, although phoneticians and their machines can distinguish between the two *t*'s, there is no context in which the replacement of one by the other changes the meaning. If a person who is not a native speaker of English uses an unaspirated *t* where he shouldn't, we feel that there is something slightly off, that he has a foreign accent in some sense, but we don't have any trouble understanding him. As far as English is concerned, the two *t*'s are "isosemantic". There is no such word—I just made it up—but everybody can probably guess what it would mean if it existed: it would mean that the replacement of one by the other preserves meaning.

What then is a phoneme? Or, better asked, what is the phoneme of a sound? Answer: the collection of all sounds isosemantic with it. Since *b* and *m* are not isosemantic, *b* does not belong to the phoneme of *m*, but the *t* in "tone" does belong to the phoneme of the *t* in "stone".

A similar analysis of the concept of species is possible, but I will not enter into it now. A dictionary might define a species as "a collection of organisms capable of interbreeding", but before we could discuss the pertinent analogue of "isosemantic", we would need to sort out the sexes, and that digression, while possibly interesting, would take too long.

The concept of number is nearer at hand, and, at least in mathematical circles, very well known. We all use words such as "five" every day, but do many people ask themselves what "five" is? And, by the way, shouldn't they be ashamed of themselves? We wouldn't use words such as "grandfather", or "tax", or "lawnmower" without being able to define them—without, to be specific, being able to tell a ten-year old child exactly what a grandfather, or a tax, or a lawnmower is, but the challenge is to tell him exactly what a number is. I don't mean what a number *does*, or how a number can be used—I mean what it *is*.

All right: what is "5"? We may not know that, but we know that if it's the answer to "How many fingers are there on your right hand", then it's also the answer to "How many players are there on a basketball team?" In other words, while we may not know what "number" is, we do know when two sets of objects (be they fingers, or whatever) are "equinumerous". They are that just when we can establish a correspondence between them (for example, by pointing to each basketball player on the team with a different finger) that is a one-to-one correspondence—each object in each set corresponds to a unique object in the other set.

What then is a number? Or, better asked, what is the number of objects in a set? Answer: the collection of all sets equinumerous with it.

Abstraction and Attitude: Equivalence Relations and Extensionalism

This is an abstract definition, it is a frightening definition, it's an ingenious definition. It is due to Bertrand Russell, and it leads me now to comment on two things: one, an abstraction, a basic mathematical concept, that includes the way species, phonemes, numbers, and many other concepts in many parts of life are best thought of, and, two, an attitude, a philosophical stand, that some mathematicians embrace, and that contributes greatly to the clarity and precision of mathematics. The name of the concept is "equivalence relation", and it is well known and standard; the name of the attitude is "extensionalism", and, while the attitude is not uncommon, the name, as far as I know, is something that I've been using for some time in private, but no one else has ever heard of.

An equivalence relation is a relation that has three properties in common with the relations of being isosemantic and equinumerous, namely that it is reflexive, symmetric, and transitive. The replacement of an utterance by itself (which is, of course, no replacement at all) surely preserves meaning, and each set has a one-to-one correspondence with itself—that's what "reflexive" means. Officially: a relation is reflexive in case every object in its realm does bear that relation to itself. So, for instance, fatherhood is not reflexive—no one can be his own father—and whether brotherhood among, let us say, human males is or is not is a small hairsplitting debate about how you want to use words. Am I my own brother?

To say that a relation is "symmetric" means that the roles of two objects in the relation can always be reversed. Example: if the initial sound in"pit" is isosemantic with the initial sound in "pendulum", then, vice versa, the initial sound in "pendulum" is isosemantic with the initial sound in "pit". Similarly: if a basketball team is equinumerous with the fingers on my right hand, then, vice versa, the fingers on my right hand are equinumerous with a basketball team. Here are a couple of non-examples: fatherhood is not symmetric (my father bears that relation to me, but I do not bear the same relation to him), and fondness is not symmetric—while it may often happen that someone I am fond of is fond of me, it is not guaranteed, and one single exception disproves the universality of the property.

"Transitivity" is just as easy a concept, but it takes a little longer to say. If three sounds are isosemantic in order, that is, the first and the second are isosemantic, and the second and the third are, then it follows that the first and the third also are.

A well-known non-example is friendship: it isn't always true that if Tom is Dick's friend and Dick is Harry's, then Tom and Harry also bear the relation of friendship to each other.

So, that's what an equivalence relation is: one that is reflexive, symmetric, and transitive. And any time we run across an equivalence relation, the objects to which it applies can be split up into what are called equivalence classes—and, using that language, I can now say that a phoneme is an equivalence class of the relation of being isosemantic, and a number is an equivalence class of being equinumerous.

That's one of my main points, and when I learned of it, I felt that I gained a thrilling insight—that's what I mean by a thrill of abstraction. The notion of equivalence relation is one of the basic building blocks out of which all mathematical thought is constructed. It is simple, it is general, it is widely applicable, and it is 100% explicit and precise. And, what's more, it has nothing to do with columns of numbers or triangles or electronic computers or whatever mathematics is sometimes thought to be about—it is abstract pure thought.

Now, about "extensionalism"—there I am not sure I can explain what I feel. In a short sentence what I am trying to say is that to a mathematician—well, in any event, to me—a concept IS its extension. Consider, for an example, the number 5. What is it? Not "What does it do?", "How can it be used?", or "How can I tell it apart from others?", but "What IS it?" Mathematicians usually ask such questions: their insistence on definitions and their insistence on complete precision in the definitions and complete consistency in their use is one of the distinguishing features of their art. The "extension" of a property (an old, established philosophical term) is the class of all objects that possess it. Thus, the extension of "blue" is the class of all blue things—the sky, the Danube, the books, the neckties, whatever —whatever—everything that happens to be blue. The extension of "5" is the class of all quintuples—basketball team, fingers of a hand, whatever. A cautious lexicographer might be willing to go this far with the mathematician: very well, he might say, 5 is the property that is common to all quintuples. The rigorous mathematician would consider that pussyfooting, however. Just what, pray, is a "property"?, he would ask. And how dare we speak of "the" common property of the set of all quintuples—how do we know there is only one? No, sir!, he would say: all I really know about fiveness is that I am willing to assert it of the fingers on my right hand, and, extending from there, of any other set that I can put in one-to-one correspondence with those fingers. In other words, he would say, I know the extension of fiveness, and that's all I know about it. The only courageous way to define 5, therefore, is to follow the principle that a concept IS its extension—and, as a religiously observant extensionalist, I therefore *define* 5 to be the equivalence class of equinumerousness to which the set of fingers on my right hand belongs.

There is something cold and forbidding, something impersonal and frightening, about this definition—one might feel that while it is intellectually, legalistically defensible, it somehow misses the essence of the concept being defined. It reminds me of the classical, and equally unsatisfying definition of a man as a "featherless biped". When I first heard that I objected. Surely, I thought, there is more to humanity than that. What about soul, what about humor, what about art, culture, technology, war, friendship, motherhood—what about all these "essential" characteristics of humanity—doesn't a cold-blooded definition such as "featherless biped" miss them all, and therefore miss the point? After many years of becoming used to the idea, I no longer feel that discomfort in the presence of an extensionalist definition. If it is indeed true (I repeat: *if* it is indeed true—I am not asserting that it is) that humanity is coextensive with the class of featherless bipeds, then humanity is the class of featherless bipeds. And, similarly, since "fiveness" jolly well is coextensive with the class of all quintuples, I happily embrace the definition according to which 5 *is* that class.

Dreamers and Non-Constructive Proofs

It's about time I turned to the second of the three questions that I raised, in order to describe a second, very different, basic mathematical belief and behavior. Is there, I asked, a number that when multiplied by itself 5 times gives the same result as adding 2 to it? There are those, both among dreamers and among very practical people, who would answer that question by yes only if they could explicitly produce a number with the property described, or, at the very least, in the worst case, if they could explicitly describe an algorithm, a procedure of calculation, that will produce such a number. Thus, for instance, if I change the number 2 in the problem to 240, and if I go on to observe that $3^5 = 243$, which is the same as $3 + 240$, then, I think, we would all agree that the changed question has been answered, and answered in the affirmative.

There is, however, another way to answer such questions, the way of non-constructive proofs, of which I'll give a modest example. Imagine that I have an ultra-efficient but not especially intelligent computer, programmed to tell me instantaneously which is greater, x^5 or $x + 2$, whenever I ask it about any particular whole number x, but that knows about whole numbers only. All right, I say to the computer, let's go: $x = 0$. It says: $x + 2$ is greater. I say: $x = 1$. It says: $x + 2$ is greater. I say: $x = 2$. It says: x^5 is greater. I say: Hurray!—the game is over, and the answer is yes. That's right, isn't it? If I imagine myself moving along the line, scanning all the numbers from 0 on up, and if I know that somewhere (say, when $x = 1$) $x + 2$ is the larger of x^5 and $x + 2$, and somewhat later (when $x = 2$) x^5 is the larger, then, by an intuitively obvious and rigorously provable property of continuity I can rest assured that somewhere in between x^5 and $x + 2$ will be exactly equal.

What do I know now that I didn't know before? Do I know a number x such that $x^5 = x + 2$? No, I don't. All I know, but that I know for sure, is that although I am not (not yet!) able to construct one, such a number does exist.

I have just given, as I promised, a modest example of a non-constructive existence proof. I call it a "modest" example, because, as a matter of fact, with a little trouble it can be converted into a concrete algorithm that will produce a number of the kind that is wanted as accurately as desired: for the benefit of the reader who is just dying of curiosity, I'll put on record that rounded off to five decimals the answer is 1.26717.

Genuine non-constructive existence proofs, the kind that cannot be converted into a computational procedure, are sometimes a source of heated debates in the mathematical family. They are impressive demonstrations of human ingenuity and of the depth of mathematical thought. Sometimes, for instance, in order to prove that a certain set (such as the set of points on the number line) contains at least one object of a particular kind (such as a number x for which $x^5 = x + 2$), a mathematician might use a "stochastic" method. That's a complicated concept whose detailed description would take us too far afield, but in qualitative terms it means something like this. Design a gambling game, a dice game, say, whose possible outcomes are the objects in the set under consideration. Using the properties that are demanded of the particular objects whose existence is in question, compute the probability that the gambling game will produce one of those objects. If that probability turns out to be a positive number (in other words, not 0), then we can be sure that the set of desired objects is not empty—objects like that must exist—even though the method of proof doesn't even yield a hope this side of heaven of ever concretely exhibiting one.

The stochastic method is a much fairer example of a non-constructive existence proof than the "modest" one based on continuity. Many non-constructive existence proofs use some notion of the "size" of a set (such as probability, or dimension, or even just cardinal number), and achieve their end by proving that the size of the set of objects not known to exist is large—large enough to guarantee that it is not zero!

Schubfachprinzip

The very first question that I asked (remember?—the handshake question?) can be used to illustrate a third basic principle of thrilling, pure, abstract mathematical thought, the so-called Schubfachprinzip, or pigeonhole principle, but I think I'll yield to my congenital tendency to mathematical sadism, and let that question stand as a puzzle for you—I'll use a different question to explain the Schubfachprinzip.

Suppose that a bunch of us are together in a room, 100 of us, say, and we form temporarily a small society of our own. In this closed society there are a certain number of acquaintanceships: some of us are acquainted with some others. I don't know which ones of us are acquainted with which others, but I'm sure of one thing: I'll bet that there are at least two of us that have the same number of acquaintances.

Believe it? Let's see if I can make it convincing. Suppose that someone asked us, each of us, myself included, "How many other people in this closed society are you acquainted with?" We could all tell him an answer, somewhere between 0 and 100. No, wait a minute. If there are exactly 100 of us, then nobody is acquainted with 100 *other* people; the largest number can be no larger than 99. As far as 0 is concerned, that's all right, there could well be some hermits among us, but it's not likely, and, in any event I can easily settle that case. If there are two or more hermits, then I've already won my bet: any two hermits have the same number of acquaintances. If there is only one hermit, then let's ostracize him—let's not count him—let's go so far as to pretend that he isn't here. I must still prove that among the remaining 99 there are two of us with the same number of acquaintances, and I'll do so—but because 100 is easier to say than 99, let me assume that even if the possible hermit is not counted there are still 100 of us left.

So then, what possible numbers will each of the 100 of us give to the questioner? Answer: any number between 1 and 99 inclusive. What is it that I am betting? Answer: that two of us will give the questioner the same number. Indeed: how could we fail? There are only 99 numbers to tell him and there are 100 of us telling: there must be at least one repetition.

Isn't that pretty? I think it is, and, by the way, it is an application of the impressive sounding but childishly easy Schubfachprinzip. The principle says that if we have a number of pigeonholes, and if there are more letters than pigeonholes, then at least one pigeonhole will end up with more than one letter in it. That childishly easy principle is still another basic building block of mathematics—it occurs over and over again, sometimes in very sophisticated contexts, and it is the backbone of all so-called finite or combinatorial mathematics.

Note that the three basic principles that I have described so far are of three different kinds. "Equivalence relation" is a concept; "non-constructive existence proof" is a technique (and an attitude); and the Schubfachprinzip is a theorem, a fact (with, to be sure, many millions of applications and very different-sounding special cases). I could have, and for greater clarity I feel sure I should have, given other examples of the domains of application of the three basic principles already mentioned, and, by the same token, I could and should have given other principles, that arise in other problems. Anything like completeness in a discussion such as this is impossible in a few pages—but perhaps I could do more justice to both the subject and the reader by at least mentioning what else could have been said.

Thus, for instance, is it obvious that the face of a clock is, in effect, a picture of an equivalence relation? (I have in mind the relation between two numbers that holds when one is obtained from the other by adding 12 to it, or, for that matter, any multiple of 12—so that the 13 o'clock is the "same" as 1 o'clock.) Or is it obvious that round-off downward (permitted by the Internal Revenue Service, or so I am informed, when we calculate our income tax) defines an equivalence relation? (In this sense a tax of $317.23 is equivalent to $317.00; more generally two possible calculated taxes are equivalent if ignoring the pennies, any number of them from 1 to 99, makes them equal.)

As for examples of non-constructive existence proofs: many of them depend on the famous (for some people infamous) law of excluded middle. Do we want to prove that a certain mathematical construct "exists"? Very well—let us assume that there is no number, or triangle, or whatever, that satisfies the definition we are working with; let us proceed to reason from that assumption, and, if we're lucky, we shall presently arrive at a contradiction. Conclusion: non-existence is untenable, and at least one instance of the object must indeed exist. This kind of non-constructive existence proof makes the people who don't believe in it angrier than most other kinds.

Do Normal Numbers Exist?

Other examples of non-constructive proofs can occur in the theory of the so-called "transcendental numbers" (there are, in the sense of Cantor's set theory, "more" transcendental numbers than non-transcendental ones, hence there must be at least one), and in the theory of "normal" numbers (the "length" of the set of normal numbers, or, in other words, the "probability" that a number be normal, is not zero, and hence there must be at least one such number).

The last thing I mentioned is sufficiently interesting that I am strongly tempted to go into a bit of technical detail. I promise it won't last long.

In this discussion the "numbers" I want to consider are the proper fractions—the positive numbers that have no whole number part, such as

$$\begin{aligned}&.5000000\ldots,\\&.3333000\ldots,\\&.3333333\ldots,\\&.142857142857\ldots,\\&.12345678901234567890\ldots.\end{aligned}$$

When we look at the decimal form of such a number, we can ask how often does the digit 8 occur in that form, in the long run average? The answer for the first three numbers is "never"—8 just isn't in the act. The answer for the fourth number is "one sixth of the time". Isn't that clear? There is exactly one 8 in each successive group of six digits; among the first million digits the number of 8's is approximately one sixth, and if we replace "million" by more and more, the approximation to one sixth becomes more and more nearly perfect. For the last number in the list, the answer is "one tenth"; the reasoning is the same as before.

There are ten digits available to us, and we might consider that a number is "fair" if it treats each of them the same as all others—in other words, if each of the ten digits occurs in that number exactly one tenth of the time, in the long run average. In this sense only the fifth of my five sample numbers is fair.

There is a more sophisticated notion of fairness, however, according to which none of my sample numbers is fair. To illustrate what I mean, let me ask this question. Given a number (in decimal form, with no whole number part), how often do the digits 5 and 7 appear in it, next to each other, in that order (in the same long run sense as before)? The answer is "never" in all my examples, except the fourth, and in that case it is "one sixth". To see what I mean, go along the digits, count all "blocks" of length two, and keep track of what proportion of them are "57".

What should the answer be if we are to regard the number as fair, fair not only to each individual digit, but fair to each conceivable pair as well? The answer depends on how many possible pairs there are—and the answer to that is 100. Clear? Sure it is: just count them, from 00, 01, 02, . . . , 09, 10, . . . , to 97, 98, 99. One hundred it is, and, consequently, the only way a number can be "pair fair" is if it has each possible pair in it one hundredth of the time (in the long run average).

Can we write down a number that is fair to each digit and to each pair of digits as well? Sure we could, with paper, pencil, and some time—but as soon as the task was finished, I'd be ready with a new question that demands to be asked. The new question is about triples, such as 293. I would now refuse to call a number fair unless it treated fairly each digit (with frequency one in ten), each pair (with frequency one in a hundred), and each triple (with frequency—surely the answer is guessable—one in a thousand). And once the pattern is clear, I can continue it: in my infinite greed for justice I'll demand an absolutely fair number, by which I mean one in which all blocks of all lengths occur with the "right" frequency (one in ten, or hundred, or thousand, or ten thousand, etc., for singles, doubles, triples, quadruples, etc.). The usual technical, mathematical name is not "absolutely fair" but "normal", and now we've got a question, a bona fide, hard, juicy mathematical question. Can all these, infinitely many, conditions be satisfied simultaneously? In other words: do there exist any normal numbers?

That one I don't think most people can do, not unless they are professional card-carrying dues-paying members of the mathematician's union. But the mathematician who is not afraid of non-constructive existence proofs, and who has a small amount of training in modern probability theory, can sail right through it. All he needs to do is to consider the process of choosing a number at random, by, for instance, throwing an arrow randomly at the segment of the number line that lies between 0 and 1, compute the probability that the number he hits is normal, and observe that the answer is not 0. The computation is not trivial—that's where some mathematical technique is really needed. The probability that it yields is not only different from 0, but it is as different as it could possibly be: it is equal to 1. In other words: it is almost certain that a randomly chosen number will be normal—which surely guarantees that normal numbers do exist.

The number of what I have called "basic mathematical principles" is surprisingly small. No one has ever listed them, and it would be a risky, controversial thing to do, but most mathematicians agree that mathematics is a unit—it all hangs together, with all subjects interwoven, and all concepts applicable everywhere—the number of bricks needed to build such a marvelously compact edifice cannot be very large.

That's one general comment; I'd like to make one more. I have been discussing the thrills of abstraction, and, in particular, the thrills of mathematics, which most people consider very abstract indeed. Would it be a contradiction if I now said that mathematics is an experimental science? I do think that mathematics is abstract, and I do think that mathematics is an experimental science, and I do not think that those two beliefs contradict one another.

To solve a mathematical problem is not a deductive act—it is a matter of guessing, of trial and error, of experiment. To solve the handshake problem, for instance, for five couples, we could do a lot worse than just plain guess. Guess, for instance, that the answer is 7, and then try it out and see what, if anything, is wrong with that guess. Another procedure, a more dignified one that more nearly deserves to be called an experiment, is to vary the conditions and try to solve some related but, we hope, easier problem. What, for instance, happens to the handshake question if we ask it for only four couples? or three? or two? or even just one?

Abstractions Are Facts

That's the sort of way that a typical working mathematician proceeds—his attitude is not that of creation but of discovery. The answer is there somewhere, and we have no control over what it is—all we are trying to do is find it. The concepts, techniques, and theorems are abstract all right—but our learning about them proceeds the same way as our learning about the boiling point of a chemical and the acceleration of a falling body. The abstractions are *facts*, facts outside of us, facts that we do not "invent" but that are there for us to find if we can.

Some readers will recognize, of course, that the position I have thereby "proved" is that of an unreconstructed die-hard Platonist, but they won't, I hope, hold that against me. My convictions (please do not call them prejudices) took a long time to grow and I would hate to have to give them up. I am convinced that mathematics is infinite in its extent and applications, yet a unity in its conceptual way of looking at things and describing them; the facts of mathematics are there waiting for us to guess at, experiment on, and finally stumble across; the concepts, the techniques, and the facts are abstract, and, in their very abstraction, one of the most thrilling phenomena of the universe.

P.S. The answer to the handshaking problem is 4.

Paul R. Halmos is Distinguished Professor of Mathematics at Indiana University, and Editor of the "American Mathematical Monthly." He received his Ph.D. from the University of Illinois and has held positions at Illinois, Syracuse, Chicago, Michigan, Hawaii, and Santa Barbara. He has published numerous books and nearly 100 articles and has been the editor of many journals and several book series. The Mathematical Association of America has given him the Chauvenet Prize and (twice) the Lester Ford award for mathematical exposition. His main mathematical interests are in measure and ergodic theory, algebraic logic, and operators on Hilbert space.