PICTURE PUZZLING

Mathematicians Are Rediscovering the Power of Pictorial Reasoning[*]

by IVAN RIVAL

MARJORIE RICE was an unlikely candidate for the role of mathematical innovator. A San Diego housewife and mother of five, she had no formal education in mathematics save a single course required for graduation from high school in 1939. Nonetheless, in 1975, she took up a problem that professional mathematicians had twice left for dead, and showed how much life was in it still.

The problem was tessellation, or tiling of the plane, which involves taking a single closed figure—a triangle, say, or a rectangle—and fitting it together with copies of itself so that a plane is covered without any gaps or overlap. A region of this plane would look rather like a jigsaw puzzle whose pieces are all identical. M. C. Escher, the Dutch artist, who, like Rice, acquired mathematical insight without formal training, was a master at tessellation, as any number of his pictures show. But the tessellations in most of those pictures involved curves. Rice worked primarily with polygons, which consist only of straight lines. More specifically, she worked with convex polygons, in which the line joining any two points on the polygon lies entirely within the polygon itself or on one of its edges. (A five-pointed star, for example, does not qualify as a convex polygon.)

By the time Rice took up tiling, its basic properties had been established. Obviously, any square can tile the plane, as many kitchen floors have demonstrated. Equilateral triangles are also a fairly clear-cut case. And there is one other *regular* polygon (a polygon whose angles, and sides, are equal) that can tile the plane: the hexagon. This fact was established by the ancient Greeks but had long before been exploited by honeybees in building their honeycombs.

And what of *irregular* polygons? As it turns out, any triangle or quadrilateral, no matter how devoid of regularity, will tile the plane. Yet no convex polygon with more than six sides can do so, and the three classes of convex hexagons that can were uncovered by the end of the First World War. So the only real question left by the time Marjorie Rice happened on the scene was, Which convex pentagons tile the plane?

The question had been taken up by Karl Reinhardt, a graduate student at the University of Frankfurt, in 1918. In his doctoral dissertation (which also settled the hexagon question once and for all), Reinhardt defined five classes of convex pentagons that tile the plane. For example, consider a pentagon whose angles are labeled A through E and whose sides are labeled a through e, in which each side is opposite the angle that shares its letter. Reinhardt showed that any pentagon in which $C+E=180$ degrees and $a=c$ can tile a plane. He did not claim that this and his four other classes of convex pentagons were the *only* such classes to be found, but he intimated as much, and this became the received wisdom; the problem of tiling the plane with convex polygons was closed.

Fifty years later, the problem was reopened and then closed again. In 1968, Richard B. Kershner, writing in *The American Mathematical Monthly*, claimed to have solved it once and for all. Kershner had found three classes of convex pentagons that could tile the plane but had been overlooked by all his predecessors. He confidently asserted that these three pentagons, together with Reinhardt's original five, constituted the final list of convex

[*] Reprinted from *The Sciences*. New York Academy of Sciences: January/February 1987, 40-46.

pentagons that tile the plane. Pleading space limitations, he did not publish a proof of the assertion.

As it turned out, space was not the only obstacle to such a proof; Kershner was wrong. In 1975, Martin Gardner wrote two of his “Mathematical Games” columns in *Scientific American* on tiling of the plane, thus prompting scores of trained and amateur mathematicians to try their hands at tessellation. Marjorie Rice was not the first to succeed, but she was by far the most prolific. Rice furtively read her son's *Scientific American* upon its arrival every month, and Gardner's December column, which reported on the progress of amateur tessellators, stirred her to action. “This was the busy Christmas season which took much of my time,” she later recalled, “but I got back to the problem whenever I could and began drawing little diagrams on my kitchen counter when no one was there, covering them up quickly if someone came by, for I didn't wish to have to explain what I was doing to anyone.”

Before long, Rice was drawing pentagons and fitting them together mentally. She soon found an undiscovered pentagon that fit snugly with copies of itself and proved to her own satisfaction that the fit must indeed be perfect for the entire class of pentagons to which it belonged. She then informed Gardner of her discovery, observing in her letter that “one of the enclosed examples in which the two sizes of line are in golden proportion makes a very pleasing arrangement, I think.” Gardner put her in touch with Doris Schattschneider, a mathematician with a professional interest in the harmony of mathematics and art. With Schattschneider's encouragement, Rice has since discovered several more kinds of pentagons that tile the plane, bringing to thirteen the number of such classes known. (The findings were reported in a 1978 article in *Mathematics Magazine,* written by Schattschneider; Rice herself has never published a word on the subject.)

The cover of the November 1985 issue of *Mathematics Magazine* featured an arresting illustration—a previously unknown tessellation of convex pentagons. Rolf Stein, of the University of Dortmund, in West Germany, had found a fourteenth class of pentagons. Unchastened by the experience of Reinhardt and Kershner, Stein has claimed that the problem is solved once and for all. This remains an unproven assertion; Marjorie Rice's final contribution to mathematics may be yet to come.

TODAY tiling the plane—or, at least, trying to—is fast becoming a minor mathematical industry. It has attracted a substantial following, especially among amateurs. Its minstrels are much in demand, servicing conferences, colloquiums, and the always hungry recreational-mathematics press. It is hard to believe, amid all the activity, that no progress was made in the subject during the fifty years between Reinhardt's doctoral dissertation and Kershner's discovery of three new convex pentagons. That hiatus, indeed, demands explanation. After all, since 1918 there had been much activity in all the sciences, including mathematics. And no technological or theoretical breakthrough was missing; tiling the plane was not a problem that required, say, high-speed computers for its solution. What, then, accounts for the period of stagnation? What did Marjorie Rice have that scores of past mathematicians did not have?

In a word, pictures. During most of this century, mathematicians have frowned upon the use of diagrams in exposition and argument. Even with a problem so unavoidably visual as tiling of the plane, proofs of solutions, preferably, would not invoke diagrams but would consist merely of rows of symbols: numerals; English, Greek, and Hebrew letters; compound characters made by stacking up bars and dots and tildes—enough symbols, all told, to give a typographer nightmares. And each row would follow from the previous row in accordance with the laws of mathematical deduction. As long as this deductivist orthodoxy held sway, there was little room in mathematical discourse for diagrams or for arguments that appealed to common sense or intuition.

Yet tiling problems call for, above all, thinking and talking in terms of images. Though it is not difficult to mount an entirely formal argument in support of the fairly simple cases analyzed by Reinhardt, the logic behind the more complex cases explored by Rice is difficult to convey without using pictures. In fact, as an aid to reasoning, she developed a symbolism all her own, a synthesis of pictures resembling hieroglyphics and an arcane code.

Rice's methodology, as it happens, mirrors a shift in the way mathematics is being done. Even as she was so vividly demonstrating the advantages of pictorial reasoning and argument, professional mathematicians were rediscovering them. This renaissance in the use of diagrams is seen throughout the mushrooming field known as combinatorics—not just in such inherently visual problems as tiling but even in problems that have no obvious connection to geometry. Row upon row of symbols is no longer the only permitted form of professional discourse. Appeals to intuition, extrapolation from one or more examples, and the use of pictures—lots of pictures—are being reinstated in the language of mathematical argument. In one branch of mathematics, at least, the tyranny of deductivism has begun to erode.

DIAGRAMS ARE, of course, as old as mathematics itself. Geometry has always relied heavily on pictures, and, for a time, other branches of mathematics did, too. Even Isaac Newton, commonly credited with the invention of the calculus, did not actually *prove* its fundamental theorems—at least, not in accordance with today's stringent standards of formal proof. In discerning the properties of classes of algebraic functions (relationships that define one variable in terms of one or more other variables, such as $y=x^2+1$), he would, we presume, select an example, contemplate its graph, draw generalizations, and test them with further experimentation. Thus, the rules he created for determining the slope of a function's graph, while useful, did not rest on a formal foundation. Had you asked him to justify them, he would likely have presented an argument that, though compelling, was loose and depended heavily on pictures.

The intuitive and often persuasive style of argument used by Newton and his contemporaries fell into disrepute during the nineteenth century, after it proved, in several celebrated cases, misleading. One case involved Newton's own calculus—specifically, the relationship between continuous and differentiable functions. Loosely speaking, a continuous function is one whose graph has no breaks, or jumps. Thus, one could draw the curve representing the function on the coordinate axes without lifting pencil from paper. A differentiable function is, in essence, a function whose curve is smooth—with no corners or peaks. Though there are many continuous functions that are not everywhere differentiable, it had been thought virtually certain that every continuous function has at least one neighborhood—even if only a very small one—in which it is differentiable. In other words, it was thought that a continuous function cannot have a corner at *every* point along its graph. The intuitive basis for this belief was simple: it is impossible to conceive of a graph that is everywhere so densely folded and precipitous.

It was thus a stunning revelation when, in 1872, the German mathematician Karl Weierstrass unveiled an example of a continuous function that was differentiable at no point at all. However improbable this idea seemed, Weierstrass had proved it rigorously, with an analysis couched strictly in algebraic symbolism. He drew no pictures of such a graph (how could he?), but the logic establishing its existence was airtight.

The same point had already been made by Bernhard Bolzano, a Czech theologian, but had fallen on deaf ears. In 1834, Bolzano discovered a continuous but nowhere differentiable function. And, in contrast to Weierstrass, he described his function not algebraically but with a sequence of pictures of ever more jagged graphs; if continued indefinitely, Bolzano argued, the sequence would yield the graph of a continuous but nowhere differentiable function.

Would the reputation of pictorial reasoning have been less tarnished had Bolzano's work become known more widely than Weierstrass's? We will never know. Weierstrass was a very powerful figure in the mathematics community, and it was his form of the argument that prevailed. He also succeeded in formalizing some basic notions in calculus—among them that of the limit—and thereby systematized the fundamental ideas of the field that was to become modern analysis. (A limit, in the case of a simple single-variable continuous function, is a value for which there is a floor or a ceiling on the vertical axis beyond which the function's value will not venture so long as the value of *x* remains within a specified range.) Today, the standard textbook description of Weierstrass is as the man who “brought rigor to analysis” and made mathematicians doubt the reliability of their mental pictures and their geometric intuition.

Meanwhile, pictorial reasoning had been called into question in geometry itself. The doubt grew around Euclid's fifth postulate—which had always been a bit suspect anyway, lacking, as it did, the simplicity of the preceding four. (The first four postulates are that two points determine a line; that a line can extend indefinitely; that a circle is determined by its center and a point on it; and that all right angles are equal.) The fifth postulate says that if two lines in a plane are intersected by a third line, and the two interior angles (the angles facing each other) on one side of the third line add up to less than 180 degrees, then the two lines will eventually intersect somewhere on that side of the third line. Anyone who thinks about this for a moment will be convinced of its truth—as, indeed, people had been for centuries. For two thousand years, no one had succeeded in proving it rigorously, but a simple sketch was sufficient to convince the skeptical. The fifth postulate was thought to be a purely logical consequence of Euclid's other postulates.

In the 1820s, a Hungarian, János Bolyai, and a Russian, Nikolai Lobachevsky, independently described geometries based on the assumption that, although the first four of Euclid's postulates are true, the fifth is false. And it was later verified that these alternative, non-Euclidean geometries, though difficult to conceive, are as internally consistent as Euclidean geometry. These non-Euclidean geometries turned out not to be mere intellectual playthings. They are well suited to describing the properties of curved space, and Einstein showed them to be central to the mathematics of relativistic physics.

THESE AND OTHER EPISODES of disorienting mathematical discovery instigated what might be called a crisis in intuition. A time-honored approach to mathematical argument—using examples and counter-examples, appealing to intuition and common sense, invoking mental pictures—suddenly seemed filled with pitfalls. Apparently, if mathematics were ever to truly possess the certainty with which it is commonly associated, all arguments would have to be cast as rigorous proofs: long chains of mathematical symbols, with each link following logically from the one before, would have to replace more tenuously connected sequences of diagrams and examples. Of course, these proofs might be so tedious and so densely encrypted that the overarching logic behind them would be lost entirely. But no matter; so long as each step in the proof accorded with the rules of logic—so long, that is, as each incremental deduction were sound—the conclusion would deserve confidence. And such confidence was a central goal of mathematics.

This deductivist orthodoxy found voice in David Hilbert, the preeminent German mathematician of the twentieth century. In 1900, at the International Congress of Mathematics, in Paris (where Hilbert introduced his famous list of twenty-three great unsolved problems), he exhorted his colleagues: “We hear within us the perpetual call: There is the problem. Seek its solution. You can find it by pure reason.”

There is obvious merit in Hilbert's position. Rigorously deductive reasoning has played an important, indeed, central role in the considerable development of mathematics during the past one hundred and fifty years. The fundamental tenet of the mathematical creed is rightly deemed by mathematician and nonmathematician alike to lie in the field's alliance with the deductive spirit. Yet, for a creed to be useful, it must tell not only the truth but the whole truth, and the deductivist orthodoxy fails to acknowledge the whole of mathematics.

There is more to mathematical discovery than the sort of tediously deductive reasoning employed in formal proofs. There are intuition, common sense, and inductive reasoning—reasoning by experimentation with examples on the uncertain but useful assumption that they are typical, in important respects, of an entire class. These routes to the discovery of a new theorem are often not represented in the series of symbolic strings that eventually serve as the theorem's official rationale. The nature of mathematical thinking is better captured by the words of Robert Musil, the early-twentieth century Austrian writer and philosopher, than by formal proofs. In his novel *The Man Without Qualities,* he compared mathematical discovery to “what happens when a dog carrying a stick in its mouth tries to get through a narrow door: it will go on turning its head left and right until the stick slips through.” Sometimes, he added, “the slipping through comes as a surprise, is something that just suddenly happens.” Proofs that fail to capture the thought processes behind them often fail as instruments of communication and, so, are useless in conveying true understanding.

The deductivist orthodoxy is also misleading in its promise of certainty. Mathematics is presented as an inexorable progression of logical deductions that together constitute a monument of eternal and immutable truth. Once coded in the formal language of deductive reasoning, an assertion is thought to be beyond doubt. But, in fact, formal deductions are sometimes so long that the tedium of writing or of reading them dulls the mind, and crucial flaws are overlooked. More than once, a “proof” long thought unshakable has crumbled under overdue scrutiny. One example is the four-color theorem, which holds, in essence, that any conceivable two-dimensional map of different nations could be colored with only four colors in such a way that no two adjacent nations would have the same color. A number of widely accepted proofs of this theorem have turned out to be flawed. Alfred Bray Kempe, a nineteenth-century London barrister, succeeded in getting admitted into the Royal Society partly on the strength of his supposed proof of the theorem. Only later was his error detected. (The theorem is itself still considered true, thanks in part to high-speed computers that have examined it on a case-by-case basis.)

Given this fallibility of even formal proofs, perhaps a revision of the notion of proof is in order. A proof, in the end, is an argument that succeeds in convincing one's peers; thus, a proof becomes one only by attaining social acceptance and remains one only by maintaining it. In this sense, the less formal arguments, invoking diagrams and appealing directly to intuition, are sometimes the most successful proofs.

THE LIMITS OF DEDUCTIVISM are at last dawning on mathematicians, thanks largely to computers. Among the core fields of theoretical computer science is combinatorics, whose problems typically involve considering various combinations and finding one that meets certain stated goals. Playing with a Rubik's Cube—twisting it this way and that in an attempt to restore each side to a single color—is an exercise in combinatorics; an astronomical number of combinations of twists is possible, but only a select few will realize the goal. Another familiar problem in combinatorics is the traveling salesman problem, in which a salesman is to visit a large number of cities and wants to find the shortest route encompassing them. Combinatorics comes into play in computer science in, for example, the design of microchips whose tiny conduits are arranged to move electrons with optimal efficiency.

These kinds of problems tend to be amenable to pictorial reasoning, which is not especially surprising, since they are by their very nature visual. But not all combinatorial problems are. Scheduling offers one example. Consider the decision facing the manager of an automobile assembly plant: the tasks involved in building a car can be performed in many different orders, and the amount of time consumed depends on the order chosen. Because of the number of parts in an automobile, solving this managerial quandary amounts to a problem in advanced mathematics. Similar scheduling problems arise in nearly all realms of human activity: preparing a multicourse dinner can be a complex matter, and space missions call for the coordination of thousands of people and their tasks.

Problems such as these, though not superficially visual in nature, are also best solved with the help of pictures. The simplest example, perhaps, is the sorting out of hotel reservations so that none conflict. Typically, a chart is drawn—a time diagram, or Gantt chart, named after Henry Gantt, the management engineer who popularized it. Each row might represent a different room and each column a different day of the year. The resultant blocks are colored in as reservations are made. It does not take a mathematician to see that this visual depiction of the problem is more practical than a formal, symbolic rendering. But in more complex scheduling problems, too, fairly simple diagrams can be of service. Thus, in another sort of time diagram, each task is represented by an arrow, with the length of the arrow corresponding to the time consumed by the task; if the tip of arrow A touches the tail of arrow B, task A must be performed before task B. This sort of diagram could help in scheduling workers on an assembly line or technicians at Mission Control.

In another realm of combinatorics, Gérard Viennot, a French mathematician, was inspired by children's Lego blocks (plastic blocks with interlocking tops and bottoms) to devise new methods of enumeration. Their applications are not confined to two or even three dimensions but involve equations whose graphic renderings are scarcely imaginable. Nonetheless, a physical model of a simpler case triggered the creation of an entire methodological structure. (The blocks he drew on paper are of much more widely varying sizes and shapes than Lego blocks, but they fit together just as snugly.) It may be a sign of the shifting winds in mathematics that the august Séminaire Bourbaki—a conservative group of elite mathematicians, based in Paris, that in the past has eschewed combinatorics and allied itself with formalism—invited Viennot to present his method three years ago.

It is not just in their thinking that combinatorialists rely heavily on diagrams. In exposition, too, they have relaxed the deductivist imperatives enough to use pictures liberally. If any single article was a harbinger of this trend, it was one that appeared fifteen years ago in the prestigious *Journal of Combinatorial Theory,* written by Jean Mayer. The article was only one page long and, apart from its title, had essentially no exposition—not, at least, in the ordinary sense—but only pictures: three diagrams, appropriately labeled. The more conventional, written exposition would have required many pages and would have made for dull reading. The pictures told all.

Combinatorics has brought, in addition to pictures, a pragmatic cast of thought that is new to mathematics. Until recently, even problems as concrete as scheduling had been treated as highly theoretical. Thus, a mathematician would be content to prove that he could, in principle, solve a complex problem; given time, he could enumerate all possible schedules, compare them, and pick the best. Increasingly, combinatorialists will not settle for such theoretical solutions. They want to actually find the optimal schedule—or, failing that, find a schedule that is assuredly close to the optimal. Thus, one task of combinatorialists is to formulate computer algorithms that will identify, within a reasonable amount of time, a schedule that, for practical purposes, is as good as optimal. This pragmatic emphasis in combinatorics is seen in its freer mode of exposition and argument and in its heavy reliance on pictures.

THERE WAS A TIME, not so long ago, when combinatorialists were the second-class citizens of mathematics. Their work, standing on its own, with little reliance on the hallowed theorems of the past, did not qualify as "real" mathematics. Worse still, their inspiration did not come from an ideal, Platonic realm, where abstract truth is the supreme value, but from a messier place—the real world; combinatorics had evolved as a means for solving recreational puzzles and problems of technology and organization in the workplace. That combinatorics research so often had immediate payoffs—in the design of efficient computer hardware or software, for example—only lowered it in the esteem of traditional mathematicians. In large part, this was because of the perverse pleasure that mathematicians have long taken in the irrelevance of their work. In *A Mathematician's Apology*, Godfrey Harold Hardy, the great British mathematician of the early twentieth century, wrote with pride that mathematics rarely has any practical application upon its creation; in this detachment from the real world lies its purity, its beauty. For most of this century, Hardy's sense of aesthetics has prevailed, and thus did combinatorics encounter resistance for much of its early life.

But, today, combinatorics has risen from its lowly origins. Though it grows out of practical questions, its theorems often have a compelling interest for mathematicians. And, because of its practical value, combinatorics is attracting a growing fraction of mathematics students. Even old and prestigious journals now publish articles on combinatorics, and there has been an explosion of new journals devoted to the subject. They are rife with pictures and with crisp, intuitively accessible arguments.

Is it possible that some of these pictures will eventually find such frequent use as to become standard symbolic terms? It would not be the first time. In geometry, for example, the signs for *angle, parallel,* and *perpendicular* derive from illustrations of the concepts they represent. The equal-sign depicts a more abstract mathematical relationship, but even it has a diagrammatic origin: at one time it stood for two lines of equal length. (It was first used, as far as we know, by the Welshman Robert Recorde, in 1557; there were then in circulation several competing symbols for equality, but once Recorde's entry was adopted by Newton and Gottfried Wilhelm Leibniz, who had independently invented the calculus, rival signs did not stand a chance.) Perhaps, one day, the lines of text in mathematics journals will be studded with diagrams that jut slightly above the tops of adjacent letters and numerals but that more than compensate in logical power and intuitive appeal for the space they consume.

***IVAN RIVAL**, at the time of publication of this article, was a professor of pure mathematics and computer science at the University of Ottawa, in Ontario, and the editor of* ORDER: A JOURNAL ON THE THEORY OF ORDERED SETS AND ITS APPLICATIONS, *which frequently publishes articles on combinatorics.*