Mathematical Entertainments

David Gale*

For the general philosophy of this section see Vol. 13, no. 1 (1991). Contributors to this column who wish an acknowledgment of their contributions should enclose a self-addressed postcard.

The Industrious Ant

Most readers of this column are no doubt familiar with John Conway's famous Game of Life. Life is an example of what have come to be called cellular automata. I will discuss here another such automaton, called the ant, and though it is very easy to describe, its behavior is interesting and somewhat mysterious.

The ant lives in the plane which is divided up into cells by a square grid. There are two kinds of cells, white and black (later we will also introduce gray cells). Initially the ant is sitting on a cell, call it the origin, heading in one of the four compass directions. It proceeds to travel from cell to cell according to the following rule: It moves one cell in the direction it is heading. When it lands on a white (black) cell it rotates

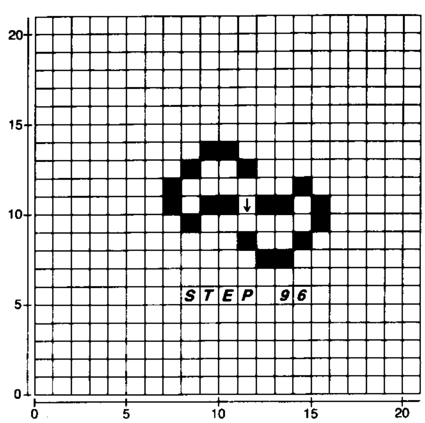


Figure 1.

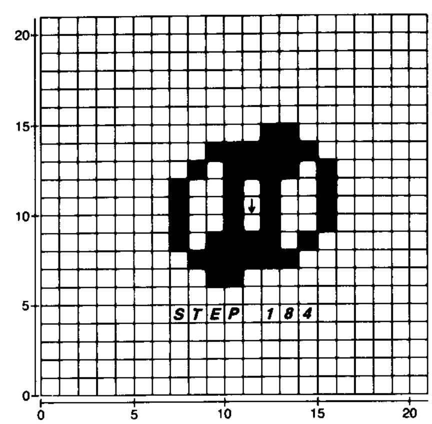


Figure 2.

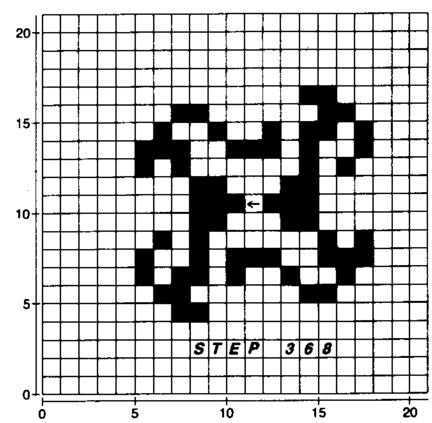


Figure 3.

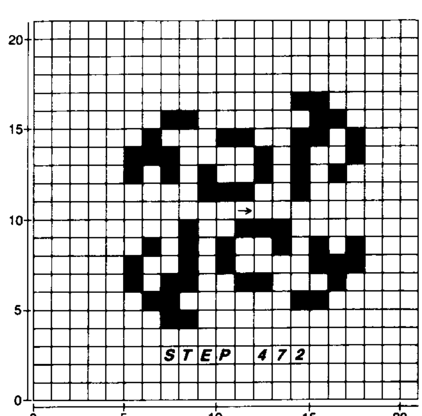


Figure 4.

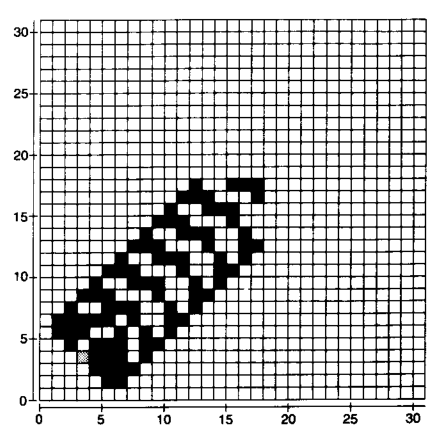


Figure 5.

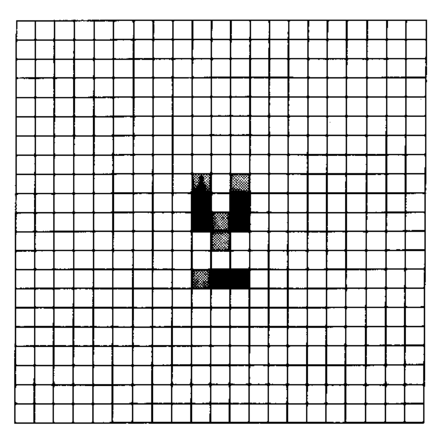


Figure 6.

* **Column editor's address:** Department of Mathematics, University of California, Berkeley, CA 94720 USA

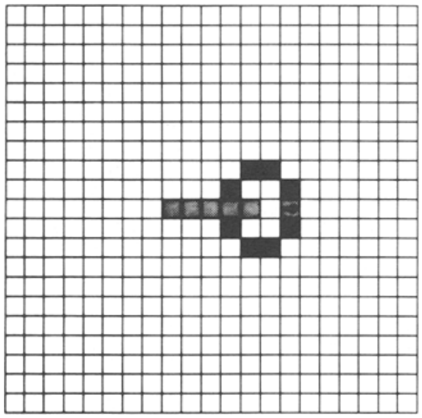


Figure 7.

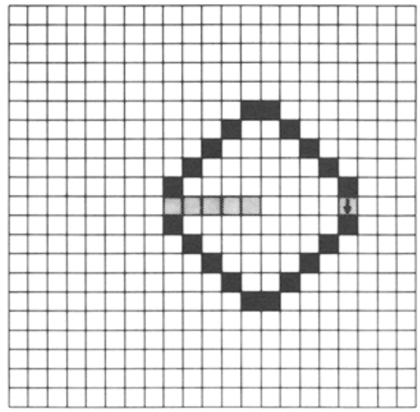


Figure 8.

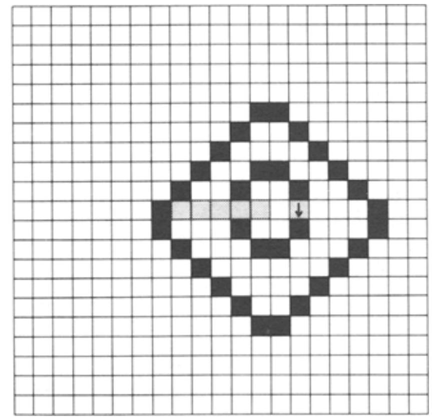


Figure 9.

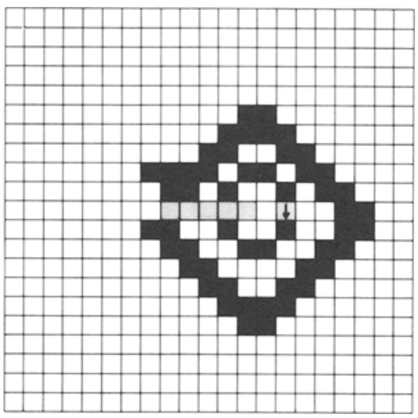


Figure 10.

its heading 90° to the right (left) and the cell then reverses its color. This is all there is to it. The game is then to start out with some given distribution of black and white cells and see how the ant behaves.

The special case where all cells are initially white and the ant is heading, say, east, is typical of what happens generally. In the early stages of its travels, say the first 500 steps, the ant, at intervals, returns to the origin, leaving behind centrally symmetric patterns of black and white cells as shown in Figures 1–4. As far as I know no one has come up with an explanation of why these patterns occur. After a while, however, things become rather chaotic for about 10,000 or so steps, but then the ant suddenly seems to make up its mind where it wants to go and heads off resolutely due southwest, leaving behind the periodic pattern shown in Figure 5, which Jim Propp, who first called attention to the phenomenon, calls a highway. On a highway, the ant takes 104 steps, ending up two units southwest of its starting point, and then repeats the process *ad infinitum*.

If the initial position includes some black cells, the ant will, of course, pursue a different course, but in hundreds of experiments it has always ended up building a diagonal highway in one of the four possible directions. Must this always happen? No one knows, but there is one quite charming result due to X.P. Kong and E.G.D. Cohen.

THEOREM. *An ant's trajectory is always unbounded.*

Proof: First note that the ant rule is reversible. The pattern of black and white cells and the ant's current position and heading determines where it came from. Thus, if a trajectory was bounded, then eventually a black–white pattern would be repeated and, hence, by the preceding observation, the path would have to be periodic, so every cell which was visited would have to be visited infinitely often. Now the key observation is that the ant's moves are alternately horizontal and vertical. This means that the cells of the plane are partitioned, checkerboard fashion, into h-cells which are always visited horizontally, from the right or left, and v-cells which are always visited vertically, from above or below. Now consider a "maximal" cell M that was visited by the ant, meaning a cell such that no cell above or to the right of it has been visited. Suppose M is an h-cell. By maximality, it must have been entered from the left and exited downward, so the cell must have been white, but then it turns black, so on the next visit from the left, the ant must go up, contradicting maximality. If M is a v-cell the argument is similar.

Neat, don't you think?

A variation on the game is to introduce a third type of cell, a gray cell, with the property that when the ant lands on such a cell it simply continues on in its current direction. Gray cells do not change their type but remain gray forever. Note that in this model the unboundedness theorem fails to hold because the partition into h- and v-cells is no longer valid, and, indeed, Cohen has found a fairly simple example of an initial configuration, shown in Figure 6, which yields a trajectory which repeats every 52 steps.

One gets some rather pretty patterns by starting with an initial line of gray cells, leaving all remaining cells white. Figures 7–10 show how, initially, the ant seems to be behaving like a spider spinning a web, but gradually asymmetries appear, and highway construction begins at about step 9000.

The ant was invented by Chris Langton who studied their behavior in some detail, including cases where several ants are moving simultaneously.

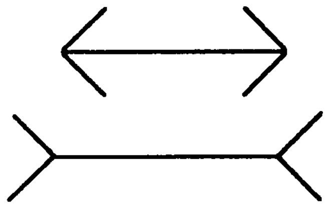


Figure 11.

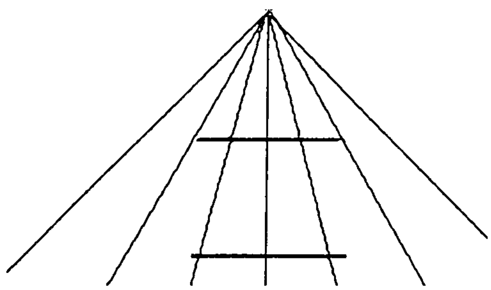


Figure 12.

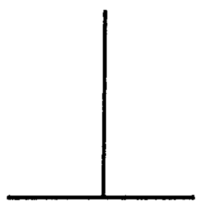


Figure 13.

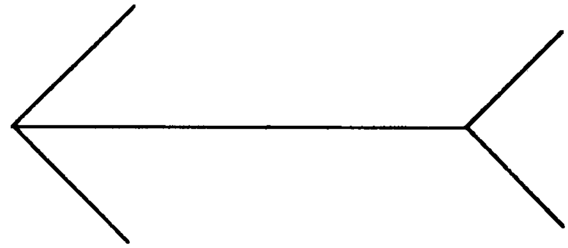


Figure 14.

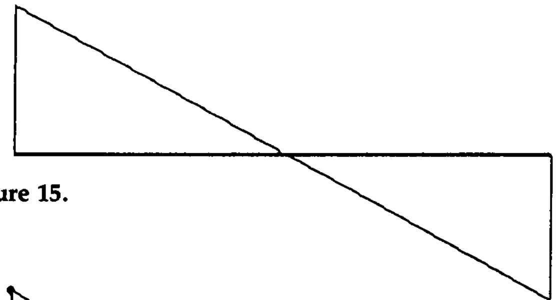


Figure 15.

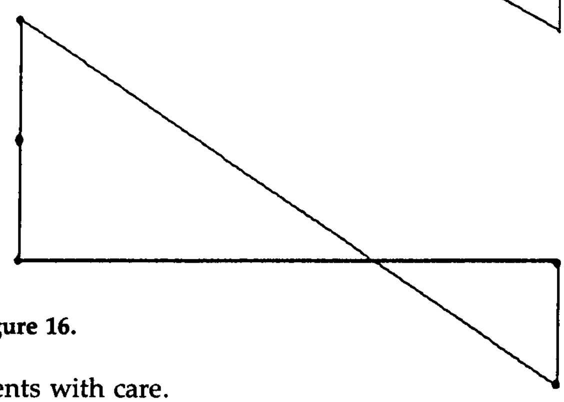
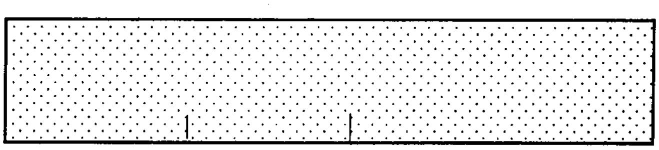


Figure 16.

Straightedge Constructions

The best way to learn a new subject, as everyone knows, is to teach it. Sometimes, if things go well, one even ends up making an original contribution. I was surprised, however, when this happened to me recently in a course I was designing for fourth- and fifth-grade schoolchildren. It was to be about geometric construction, not with the traditional ruler and compass but with a "markable" straightedge. That is, the only equipment supplied was to be a pencil, an eraser, and a rectangular strip of white cardboard on which one could make (and erase) marks.



By way of a warm-up, the students were to be asked to use this equipment to determine which of the pair of segments in the Figures 11–13 was longer. As a variation, they could then be asked to make a dot at the point which they consider to be the midpoint of the arrow shaft in Figure 14, and then use the straightedge to find out whether their guess was to the left or right of the true midpoint. The idea here is to see if the students are able to use the straightedge suitably marked to compare the distance of their guessed point from the two ends of the arrow. The well-known optical illusions used here are intended to convince them that there is some point in making these measurements with care.

This last exercise leads to the first example of a construction problem. Given a line segment, how can we use the straightedge to find its true midpoint? My proposed solution, which will probably have to be shown to the students, is to construct perpendiculars of equal length at the two end points of the segment and connect them as shown in Figure 15. (I am assuming at this point that the straightedge is a true rectangle so that constructing perpendiculars is immediate. I will return to this later.) This can lead to some significant discussion. One asks the students to explain how they know that this construction really does find the midpoint, which leads into the important subject of symmetry. (As an interesting digression, one may change the rules and allow the students to use an ordinary sheet of paper along with their pencils to find the midpoint and see how many of them come up with the idea of marking and folding the sheet appropriately. In my limited experience a child is just as likely as an adult to figure this out.)

The next step is to divide a segment in 3 equal parts using the construction in Figure 16. Again there is the opportunity for some pertinent discussion as to why one believes the length of the right-hand segment is one-third of the total length. After this, the students could be asked to divide a segment into 5 equal parts on their own and so on.

Figure 17.

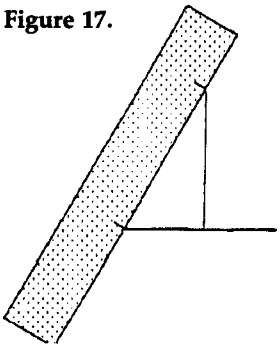
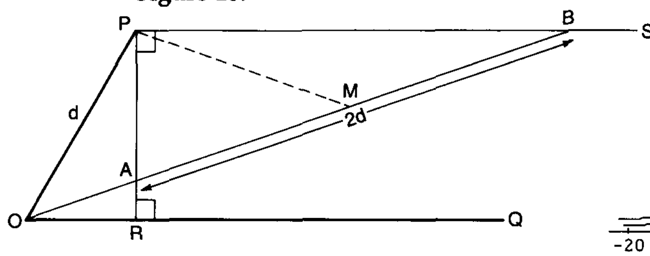


Figure 18.



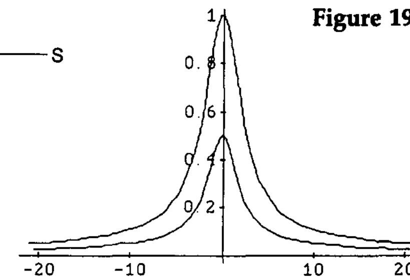


Figure 19.

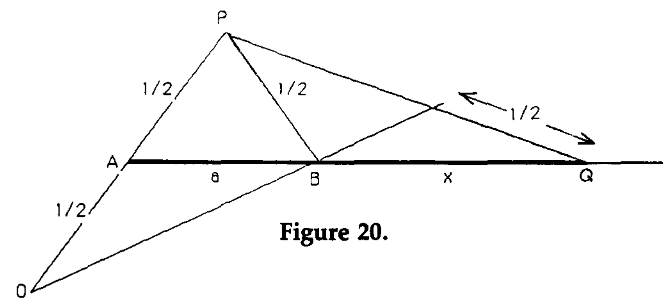


Figure 20.

One can now do a variety of things. A nice exercise, for example, is to give the students a horizontal line segment on a sheet of paper and ask them to construct the equilateral triangle having this segment as a base. The idea is to construct the perpendicular bisector of the segment and then mark the end points of the segment on the straightedge (Fig. 17).

It is now necessary to be a bit more explicit about which operations are permitted with the marked straightedge. As illustrated by this last example, one is able to find the intersection of a given line with a circle of given center and given radius, even though one is not able to draw the circle. Otherwise stated, given a line L and a point P not on L and given two marks, A and B, on the straightedge, one may obtain a new point Q on L by placing the straightedge so that A falls on P and B lies on L, assuming, of course, that this is possible. Indeed, all the classical ruler-and-compass constructions are possible because it is not hard to see how one can construct "square roots of segments."

The next exercise was to show how to bisect an angle (by making marks on each ray at equal distances from the vertex and constructing the perpendiculars at these points and connecting their intersection to the vertex). I was about to write some remarks to the effect that although, as was shown earlier, we can "trisect" segments with our straightedge, the analogous construction for angles has been proved to be impossible, and then I realized that this assertion was false! Herewith a trisection, attributed to Pappus, which uses only a marked straightedge (Figure 18).

The angle BOQ is one-third of POQ.

Proof: As indicated, the length of segment AB is twice that of OP. Now draw the line from P to the midpoint M of AB. Then the length of PM is d because—and you can take it from there.

So the question is, what is it that the marked straightedge can do that a ruler and compass cannot? The construction above proceeds by choosing the point P arbitrarily and then dropping the perpendicular from P onto OQ. Now one makes marks, X and Y, $2d$ units apart on the straightedge, and one must place the straightedge so that (1) it passes through the point O, (2) the mark X lies on PR, and (3) the mark Y lies on PS. This is the crucial maneuver. More generally, one is given a line L and a point P not on L and two marks X and Y on the straightedge. Now consider the locus traced out by Y when the straightedge is placed so that it passes through P and the point X lies on L. This locus, known as the conchoid of Nichomedes, is given by a 4th-degree equation. Typical graphs are shown in Figure 19. Of course, one is not allowed to draw the conchoid, but one can find its intersection with a line, as in the Pappus construction. Manually, this amounts to placing the straightedge so that two given marks, X and Y, lie on a pair of given lines, and then "sliding" the straightedge, keeping the marks on the lines, until it passes through a given point, a fairly easy operation to perform.

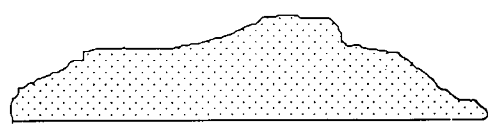
Now the interesting fact is that this conchoid maneuver allows one not only to trisect angles but, in fact, to find the roots, real or complex, of any polynomial of degree at most 4. The trick is to show that it is possible to construct cube roots of positive numbers. Figure 20 shows how to do it. The claim is that **x** is the cube root of **a** (boldfaced letters are used here for lengths). This can, of course, be proved analytically, although the calculations can get rather tangled if one does not set things up in a convenient manner. There is a slick proof making use of the theorem of Menelaus.

In making the construction, one chooses the unit of length, in this case OP. Now lay off the segment AB of length **a**, construct the isosceles triangle APB, the point O, and the prolongations of the segments AB and OB. With marks at distance ½ apart on the straightedge, one then uses these two lines together with the point P to locate the point Q by means of the "conchoid maneuver."

Finally, using trisections and cube roots, one is able to take cube roots of complex numbers (de Moivre's theorem), and these together with square roots are sufficient to find the roots of any polynomial of degree at most 4.

The point of all this is the observation mentioned earlier that using this seemingly primitive tool one can make constructions which are impossible with the classical ruler and compass. Thus, for example, from Galois theory one sees that it is possible to construct a regular heptagon as well as a 13-gon and a 19-gon, but not an 11-gon. In general one can construct a p-gon for those primes p such that $p - 1$ has only 2 and 3 as prime factors.

Returning now to the subject of constructing perpendiculars, one may ask whether the straightedge needs to be equipped with built-in right angles. Suppose, for example, it looked like this:



Can one using only such a straightedge erect perpendiculars? It turns out one can. Namely, to erect a perpendicular to a given line at a given point, . . . but on second thought, let me leave this construction as an exercise. My solution involves drawing two auxiliary lines and two auxiliary points, and making two marks on the straightedge. On the other hand, to drop a perpendicular from a given point to a given line costs me four lines and three straightedge marks, but no additional points. As a further diversion, suppose one is to draw a line through two points A and B whose distance apart is greater than the length of the straightedge. I will assume that by taking careful aim one can draw a line from A which misses B by no more than the length of the straightedge. I then have a somewhat cumbersome construction; I would welcome and publish any tidier ones.

Needless to say, none of the foregoing results is new, with the possible exception of the very elementary facts of the preceding paragraph. One source for the material is an interesting book by the Danish geometer J. Hjelmslev called *Geometriske Eksperimenter* which has been translated into German as a Beihefte to the *Zeitschrift für Mathematischen und Naturwissenschaftlichen Unterricht*, 1915. According to Hjelmslev, the cube-root construction given above appears without proof in Newton's *Arithmetica Universalis* (Cambridge, 1707), but Hjelmslev says Newton's proof is only a slight variation on one given by Nichomedes.

I suppose one could carry this further and consider constructions which can be made using both a marked straightedge and a compass. This would produce loci of degree at most 8. Perhaps such studies have been made. Any information on the subject will be gratefully received and subsequently reported.

Solutions

Another cheerful fact about the square of the hypotenuse (91-7) by D. J. Newman

Let M be a finite set of points in a right triangle. Prove that there is a polygonal path through the points such that the sum of the squares of its segments does not exceed the square of the hypotenuse.

Solution by Klaus Lagally, Germany (slightly rephrased)

We consider polygonal paths P, whose vertices are points of M. Let us denote by $k(P)$ the sum of the squares of the lengths of its segments. If M is a finite set in a right triangle, we call a polygonal path P *admissible* if it connects the endpoints of the hypotenuse, if P has as other vertices the points of M, and $k(P)$ is at most the square of the hypotenuse. The theorem clearly follows if we can prove the existence of admissible paths, since deleting the first and last segments if necessary reduces $k(P)$.

Lemma. *Let ΔAOB be a right triangle with right angle at O, and let OD be its altitude. If there exist admissible paths P in ΔADO and Q in ΔODB, then there exists an admissible path in ΔAOB.*

Proof: By admissibility, $k(P)$ is at most $(AO)^2$ and $k(Q)$ is at most $(OB)^2$, so $k(PQ)$ is at most $(AB)^2$ by the Pythagorean theorem (where PQ stands for concatenation of paths). The inequality will still hold if this path is shortened by replacing the segment from P to O and the segment from O to Q by the segment connecting the appropriate vertices directly, by the law of cosines, because the angle between the segments at O cannot be obtuse. This gives the desired path.

Proof of theorem by induction on the number of points of M. For one point the result follows, again by the law of cosines, since the angle at that point cannot be acute. When there are n points, there are two cases.

Case I. There are points of M in both ΔADO and ΔODB. Then, by the induction hypothesis, there are admissible paths in both triangles, and the result follows from the lemma.

Case II. M lies in, say, ΔADO. Then draw the altitude from D to AQ, and repeat this process if necessary until Case I occurs (as it must before the diameters of the triangles become smaller than the diameter of M). Repeated application of the lemma then gives the result.

Note that if any point of M is in the interior of the triangle, then the inequality is strict.