Cover Story

RANK-AND-FILE THINKING

BY BRIAN HAYES

You stride from cell to cell across a spreadsheet landscape, formatting A5, adjusting the formula in C26, adding to a chain of macros in Z1..Z8, then paging down to C100 for a peek at the bottom line. It is all done with a flick of the wrist or a tap of a mouse. You know where you are, where you want to go, and how to get there. The terrain is familiar; the navigation, second nature.

What makes it so easy? Why do we find the electronic spreadsheet so intuitive and straightforward? After all, a newcomer to such a program has much to master: the syntax of a macro command language, the distinction between a numeric label and a numerical value, the curious effects of circular dependencies among formulas, ranges, absolute and relative references, and more. The essential idea, however—the underlying scheme of cells organized into columns and rows—seems immediately and spontaneously comprehensible. The basic operating rules of a spreadsheet seem to be as obvious as those of a hammer or a shoe. Why?

In the case of the hammer and the shoe, anatomy explains all: They fit the hand and the foot. A similar, although more speculative, explanation of the spreadsheet's ready acceptance can also be proposed: Perhaps spreadsheets are peculiarly well fitted to the human mind. Such a hypothesis is unlikely ever to be proved or disproved, yet there is much circumstantial evidence in its favor. Gridlike patterns of intersecting columns and rows are exceedingly rare in nature, but they are quite common in human culture. Indeed, they turn up often enough, and in contexts unlikely enough, to suggest a human predisposition to organize the world in two-dimensional, tabular form.

A good place to begin is with geography, with the lacework of parallels and meridians in which cartographers have enmeshed the globe. Strictly speaking, of course, lines of latitude and longitude do not form a grid, but they represent the closest approximation possible on a spherical surface. Furthermore, away from the poles and over small distances, lines of latitude and longitude do mark off the world into rectangles much like the cells of a spreadsheet. The resemblance is even clearer in the case of a road map, where positions are identified by the intersection of a lettered column and a numbered row. With latitude and longitude it is the line that is labeled. On the road map it is the box or cell bounded by the lines that supplies the information.

Nature does not suggest these grids; we impose them

KINUKO Y CRAFT

Why do we organize our work and our world into neat two-dimensional boxes, even when logic and experience dictate otherwise?

The answer may lie with the law of gravity.

Brian Hayes writes on a variety of topics, scientific and otherwise. He lives at 45° north latitude and 93° west longitude.

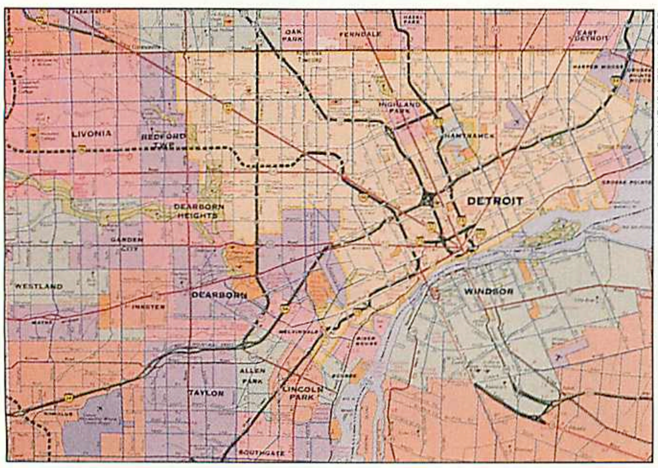
NATURE DOES NOT SUGGEST GRIDS: WE IMPOSE THEM ON THE LANDSCAPE FOR OUR CONVENIENCE.

on the landscape for our convenience. Where we physically shape the landscape, abstract grids often become concrete. Midtown Manhattan provides a well-known example: Numbered avenues go north and south and numbered streets go east and west. Without map or compass, visitors can readily find their way from First Avenue and 48th Street to Fifth Avenue and 59th, provided that they are not led astray by the intrusion of Lexington, Park, and Madison avenues. (The real world is rarely as tidy as our mathematical models.) Other cities name their streets in an alphabetical sequence or according to the chronology of American presidents. These city plans bring order to the world in two ways. The grid itself provides a rectilinear frame of reference in which all turns are right angles, while the ordering of the street names according to a set of rules provides a means of measuring relative distance and establishing position. The landscape thus becomes a two-dimensional Cartesian plane, with perpendicular X and Y axes.

The same Cartesian coordinate system is used on smaller scales. Consider the labeling of airplane seats. Boarding a commercial jetliner is often a hectic and frustrating exercise, but suppose the seats were not numbered from fore to aft or lettered from port to starboard. Suppose they were given simple sequential numbers.

LEW LAUSE/UNIPHOTO

Latitude and longitude lines. Strictly speaking, the system is not a grid, but it represents the closest approximation possible on a spherical surface.



JOHN NEUBAUER/UNIPHOTO

A road map. Positions are identified by the intersection of a lettered column and a numbered row.

ARTHUR D'ARAZIEN/IMAGE BANK

A concrete example of an abstract grid: Midtown Manhattan, with numbered avenues going north and south and numbered streets going east and west.

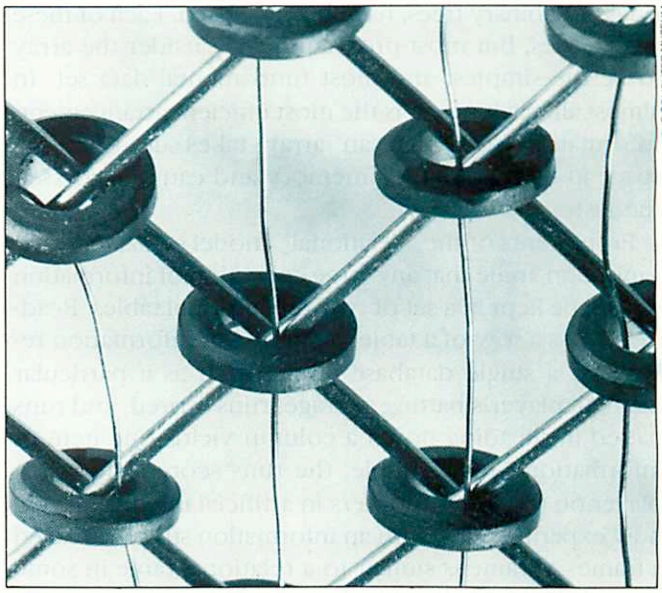
IF YOUR TICKET STUB READ "SEAT 38,988," YOU MIGHT NOT FIND YOUR PLACE UNTIL THE SECOND INNING.

Consider the same situation at a theater or a stadium. Now, with a little help from the ushers, you can find 38J Seat 3 at Yankee Stadium fairly quickly. If your ticket stub read "Seat 38,988," you might not find your place until the second inning.

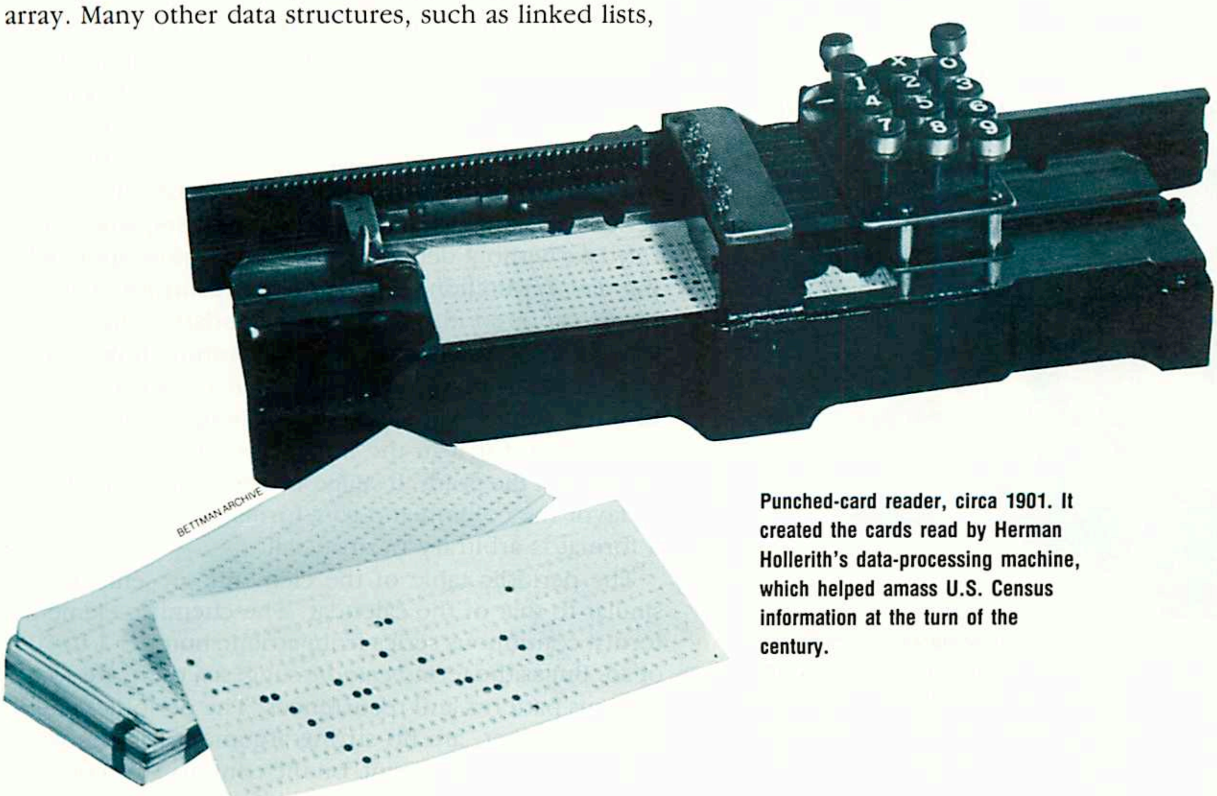
THE COMPUTER AS METAPHOR

The computer on which a spreadsheet program runs supplies many more examples of tabular organization. A typical screen displays 2,000 characters. If you write a character at position 577, where will it appear on the screen? The position is appreciably easier to visualize if it is expressed as row 8 and column 17; many programming languages allow the screen to be addressed in just this manner. Each of the memory chips within the computer is a square matrix of cells, and each cell is identified by a column-and-row address. The chips in turn are mounted in a rectangular arrangement. In the magnetic-core memories of earlier digital computers, the lattice structure was even more conspicuous. Doughnut-shaped magnetic cores were strung on woof-and-warp threads in a rectilinear mesh of fine conductors.

In computer software the column-and-row format corresponds most closely to the data structure called an array. Many other data structures, such as linked lists,


THE COMPUTER MUSEUM

Ferrite core memory. Used in earlier digital computers, this memory-storage system consisted of doughnut-shaped magnetic cores strung on woof-and-warp threads, each core able to store one bit of information.


BETTMAN ARCHIVE

Punched-card reader, circa 1901. It created the cards read by Herman Hollerith's data-processing machine, which helped amass U.S. Census information at the turn of the century.

AN APT SPREADSHEET METAPHOR IS FOUND IN A RETANGULAR ARRAY OF COMPUTERS, ALL OF WHICH ARE OPERATING IN SYNCHRONY.

rings, and binary trees, have been devised. Each of these has its uses, but most programmers consider the array to be the simplest and most fundamental data set. In almost all machines, it is the most efficient arrangement. Information stored in an array takes up the least space in the computer's memory and can be accessed the fastest.

Proponents of the "relational" model of database organization argue that any large collection of information should be kept in a set of two-dimensional tables. Reading across a row of a table yields all the information related to a single database entry, such as a particular baseball player's batting average, runs scored, and runs batted in. Reading down a column yields one item of information—for example, the runs scored—for each player on the team. Workers in artificial intelligence are now experimenting with an information structure called a frame. A frame is similar to a relational table in some respects but more flexible. Frames are two-dimensional, tabular representations of human knowledge that describe the attributes of an object. Each entry in this spreadsheetlike grid provides information on a specific aspect of an object. The developers of frames have a lofty goal. They hope to devise a way of representing all knowledge about the nature of things in an ever-expanding hierarchy of two-dimensional frames.

Entire computers can be arranged in a matrix. One of the first experiments in parallel, or multiple-processor, computing was the Illiac IV, a confederation

THE COMPUTER MUSEUM

Illiac IV, an "array-processing" computer designed by the University of Illinois and Burroughs in the late 1960s. The computer consisted of 64 computers wired into an 8-by-8 matrix. Computations were performed in parallel.

of 64 computers wired in an 8-by-8 matrix. Plans are now afoot to build much larger parallel machines, with thousands of processors, and again, in some cases, they would be connected in a rectilinear array. One reason for favoring this arrangement is that many of the problems such machines might solve lend themselves naturally to a lattice representation, although often a lattice of three or four dimensions rather than two. For example, weather can be simulated by having each computer in a three-dimensional lattice monitor conditions within a cubical volume of the atmosphere.

On the subject of parallel computation, it is worth noting that an apt metaphor for a spreadsheet program is found in a rectangular array of computers operating in synchrony. Each cell of the spreadsheet is imagined to have its own computer, which continually evaluates whatever formula has been entered there. If a cell is empty or contains a constant, the evaluation is a trivial operation. A machine that actually had a separate computer for every cell would exhibit extremely high performance when running a spreadsheet program. That's assuming, of course, that adequate quantities of data and instructions could be supplied to such a voracious computer.

FORCING THE MATRIX

Perhaps the most intriguing tabular patterns are those in which the matrix is not an intrinsic element of the structure but is imposed from outside. A good example is the calendar. We usually think of time as a linear sequence of events or perhaps as a cyclic repetition of months and years. As a basis for a calendar, the infinite series seems to appeal only to astronomers, who count the days without demarcation (they are now approaching day 2.5 million). The sense of recurrent cycles is nicely captured in the circular calendars of the Maya. The common calendar in Western culture, however, is a two-dimensional, rectangular matrix. We evidently find it a convenient shape, even though there is little justification for it in the information being recorded. That is just the point: It suggests some human tendency to favor the column-and-row format, even when such a format is arbitrary or artificial.

The periodic table of the elements presents a case similar to that of the calendar. The chemical elements form a continuous series from atomic number 1 to 100 or so. But as the term *periodic* suggests, there are some obvious patterns and recurrences. The periodic repetitions, however, are no strong argument for choosing a rectangular arrangement; on the contrary, the conventional table works only after making several ad hoc ad-

A GRID—EVEN WHEN IMPERFECT—PUTS THINGS IN THEIR PLACES AS NO OTHER ARRANGEMENT CAN.

justments. There are wide gaps in the first three periods but long insertions in the last two. Elegant spiral tables have been developed that eliminate most of these fixes; the shape accommodates quite naturally the increasing number of elements in the later (heavier) periods. Nevertheless, the rectangular table remains in almost universal use. In part this may be because of familiarity, but perhaps there is a deeper reason. While the spiral is elegant and enlightening, it is difficult to find what you need to know. A grid—even when forced and imperfect—puts things in their place the way no other arrangement can.

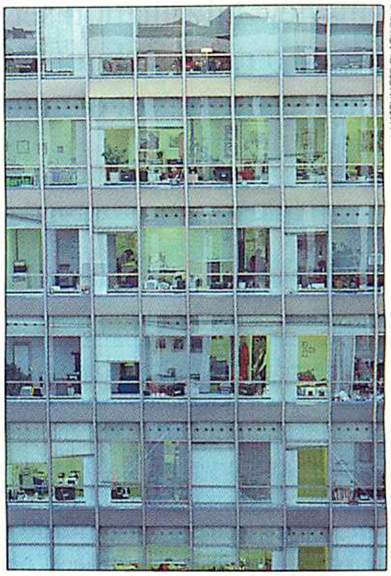
Once you start looking for column-and-row formats, they turn up everywhere. There are tables of the temperature-humidity index, of the wind-chill factor, and of optimum weight as a function of height and age. There are railroad timetables and Chinese restaurant menus. We are all too familiar with the lengthy and exacting tables supplied annually by the Internal Revenue Service. We play games from chess to tic-tac-toe to Scrabble on a matrix of squares. And of particular relevance is the grid that served as the model for the first electronic spreadsheet: the accountant's ledger, which may ultimately have its roots in some Sumerian merchant's clay spreadsheet.

WHENCE OUR TWO-DIMENSIONAL PREFERENCE

For the sake of argument, suppose this hypothesis is correct and there is an unconscious tendency in the human mind to perceive the world in two-dimensional cubbyholes. What is the origin of that tendency? It may go back to geography, to the question "Where am I?" We are gravity-bound creatures of the land, of a two-dimensional medium. We are accustomed to defining our position in terms of two spatial coordinates. One dimension provides too little structure, and a third dimension proves largely irrelevant. On the Cartesian plane we travel across in a spreadsheet program such as Lotus *1-2-3*, navigation is a visual faculty, not a cognitive one. We know without thinking that I21 is just a Page Down and a tab away from Home.

A preference for a planar map with Cartesian coordinates need not be inborn or hard-wired. A culturally acquired preference is no less real. Perhaps the origin of the column-and-row mentality will become apparent in some distant space colony. Here, humankind may raise its first generation of children, children who will learn to soar as earlier infants learned to crawl. To them, a three-dimensional spreadspace may provide the most apt electronic metaphor for the world in which they live. ■

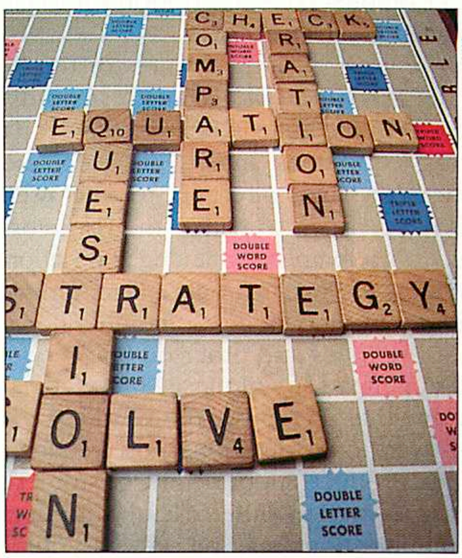
Glass-wall architecture reveals the underlying office building grid.



I. MATSUMOTO/BLACK STAR

Rank-and-file thinking at its purest: the game of chess. Modern notation for chess games uses letters and numbers to represent ranks and files on the chessboard.

LEW LAUSE/UNIPHOTO



UNIPHOTO

The game of Scrabble. Players vie to create lattice structures of words assembled from letter tiles.