Spreadsheets

W. Brett McKenzie
Roger Williams University, USA

INTRODUCTION

The spreadsheet and the personal computer are intrinsically bound. In the early 1980s, the spreadsheet was the first "killer app"—the software application that drove people to buy a personal computer. While specialized computers for word processing were replacing typewriters and database software captured the computing power of mainframes, the spreadsheet allowed managers to track, analyze, and model decisions, especially financial decisions, using a tool with low barriers to entry on an affordable computer.

HISTORY

Dan Bricklin, while a student at Harvard Business School, conceived the electronic spreadsheet and developed a simple prototype in 1978. Bob Franskton, a friend from MIT, assisted with the programming to develop the Visicalc, the first commercially available spreadsheet for the personal computer. Originally programmed for the Apple][, Visicalc was ported to a number of micro-computers as well as handheld calculators (Bricklin, undated).

Mitch Kapor, originally a product manager for Visicalc, developed the Lotus 1-2-3 spreadsheet in the early 1980s. Capitalizing on the introduction of the IBM PC, Lotus became one of the best selling software products ever published. The combination of the IBM PC with the Lotus 1-2-3 spreadsheet, transformed the PC from a hobbyist's toy into an important business tool (Power, 2004).

Just as the developers of Visicalc had not anticipated the role of the IBM PC, so did Lotus miss the importance of the graphical user interface and its instantiation as Microsoft Windows. Microsoft had developed the Excel spreadsheet for the Macintosh computer and used its experience there to produce Excel for Windows, the subsequent dominant spreadsheet with a market share in early 2005 of 90% (Liebowitz & Margolis, 2001).

Excel's most recent challenge is coming from Google, which has developed an online spreadsheet to complement its other online applications. While Google's spreadsheet does not contain as many features as Excel, it has the notable ability to leverage the communicative aspects of the Web by allowing for interactive sharing (Google, 2006).

It is suggested that the electronic spreadsheet originated with the work of Mattessich in the early 1960s (Mattessich, 1961, 1961-1964). However, the Bricklin/Franskton concept, which is the current model, required interactive technologies, such as a mouse, that were not yet developed. Furthermore, the Bricklin/Frankston concept is grounded as a more general purpose tool than the Mattessich model, whose domain was accounting.

CONCEPT

The spreadsheet has become so much at hand that the subtleties of its design, like those of the paper clip, are simply taken for granted. Furthermore, the spreadsheet has matured along with the personal computer, such that original concepts, for example, using a mouse as a pointing device rather than arrow keys, have become integrated as a natural evolution. Furthermore, earlier problems, such as the speed of recalculation, have become secondary, reinforcing the Bricklin/Franskton model that required automatic recalculation. The following discussion intertwines a conceptual perspective and an instantiation using Microsoft Excel as the example application. For the purposes of the discussion, the spreadsheet is a general personal computer application, which runs in a graphical environment. Specialized numerical software, either in statistics or accounting that uses a similar tabular layout or text-based finance programs, fall outside the definition for this article.

Visual Operation and Presentation

The current spreadsheet is a multi-dimensional, addressable, ordered array of cells whose contents may be text, values, formulas, or functions, which is able to display and store data and evaluate expressions. The spreadsheet presentation is made up of rows and columns, where the rows are usually identified numerically and the columns are identified alphabetically. A cell, the intersection of a row and column, is therefore addressed by its column heading then row location. For example, the cell in the upper left, or first cell, is A1—indicating the first row and first column. The address at the 29th row and 27th column would be AC27, indicating the 29th column by repeating the first letter of the alphabet followed by the 3rd letter. The current version of Microsoft Excel can create a sheet of 256 columns by 65,536 rows, with the address of the last cell being IV65536.

Operating with a spreadsheet is visual, spatial, and interactive, primarily by using a mouse or cursor keys to select cells, although it may be done programmatically via a scripting language. Once a cell is selected, data may be entered. This data may be text, values, formulas, or functions, or a combination thereof. Text includes the use of alphanumeric data, such as "Quarter 1" to indicate the first quarter, or numeric data, such as a Social Security Number or SKU, whose purpose is purely identification. Values are numeric entries which may be used in mathematical operations. Formulas are equations written by the user that the spreadsheet evaluates. Functions are complex formulas programmed into the spreadsheet, such as the PMT, or payment function, in Microsoft Excel, which calculates the monthly payment on a loan at a fixed rate. The requirement that the application recognizes the data type of the cell entry is integral to the spreadsheet. Excel implements this by aligning text to the left of the cell and values to the right.

A single spreadsheet can be thought of as having multiple views. The most common view is the presentation view. In its simplest terms, this view can display or hide grid lines or introduce visual features to cue the user, such as presenting negative numbers in red, as is the accounting convention. More importantly, the format of the cell may cause the underlying value to display differently. For example, a cell with the number 0.0495 formatted as currency will display $0.05. Similarly, a date, represented by a numerical value which begins with the value 1 representing January 1, 1904, may show the data with the month as text, for example "Jan 01 1904", or numerically, "01/01/1904". Despite the presentation of the value in the cell, any calculation, however, will use the value stored in that cell. The result, therefore, of 0.0495 * 20 will be $0.99 and not $1.00 were the operation to have been 0.05 * 20. The storage of value data can lead to errors by those unfamiliar with the accommodations computers make for binary storage and floating point operations.

The presentation view will obscure formulas or functions that may be present in the cells because it shows only the results of such operations. It is possible to switch the view to the formula view that shows the mathematical operations that produce the results. One may also see the actual cell contents in an area independent of the tabular sheet where the formula may be typed, commonly known as the formula bar.

Individual spreadsheets are often gathered into a workbook of multiple sheets. These sheets may be interlinked so that cell contents may be referenced by other sheets. This is known colloquially as a 3-D reference—the first two dimensions are on the sheet, with column and row identification, while the third dimension is a subsequent sheet. It is, in theory, possible to address any cell in any spreadsheet if the full path, including the machine name, directory, filename, and cell location, are known.

The Power of Cell References

Mathematical expressions or formulas, such as adding two numbers, or functions, such as computing the monthly payment, may use specific values, for example 2 + 3 or an interest rate of 5%. However, the power of the spreadsheet comes from creating expressions using cell references, such as A1 + A2. Four reasons drive this power.

First, the result is automatically recalculated. Bricklin reports that the automatic recalculation was inspired by watching his professor erase and re-enter changed values in an example on the blackboard (Bricklin, undated). It is the ability to instantly recalculate that allows a spreadsheet to perform "what if" analysis where the result may be seen when values change. The concept of referencing cells also encourages the development of a spreadsheet for a general solution rather