

Condensed Display of Multivariable Data

George C. Moon, M.Sc., P.Eng

<http://moonlite.office-on-the.net/~moonlite/Moonlite>

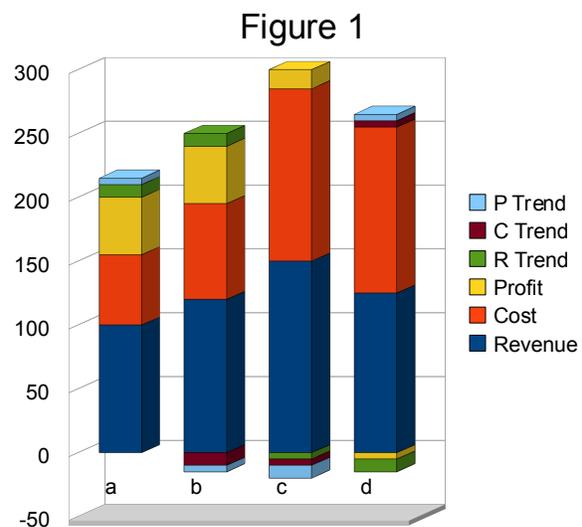
Introduction

There are a number of ways to display data for multiple variables to users. Spreadsheets and business intelligence software routinely handle multiple variables. Depending on the complexity of what is being displayed, simple tables or the use of “pivot” or “datapilot” tables are often used to present the data. Business intelligence software have additional sets of techniques to analyze and handle multivariable data.

Spreadsheets can be difficult to see, at a glance, information of importance to the viewer. Filtering can be done to find data of interest or formula can be created to highlight specific cells. Summary data can be created to reveal results. Nonetheless, if the data sets are very large, information remains hidden to the viewer. Graphics in the form of charts can draw the user to information of interest. Making quick sense visually for multiple variables can still be challenging however using traditional charting methods like pie or bar charts.

Let's take a simple example of revenue, cost and profit for a set of retail stores and assume we also desire to see trending information. That is, are revenue, cost and profit increasing or decreasing over a period of time. Figure 1 shows a subset of stores in bar chart form that tries to display the information in a compact graphic.

The relative magnitudes of the results can be seen and you can get an idea of which store shown is performing well but with some effort. Store c for example has the most revenue and is making profit, however looking at the light blue the profit is declining downward as is revenue. Store b shows that it has what looks like good profit but the trend is for costs to be worse and profit to be less even as revenue increases. An alternative way to view



information for c is  where you easily see

revenue is rising with a strong upward trend (not obvious from the chart), the profit in the acceptable range (not clear from the chart) however it is trending down (thus the yellow caution colour) and showing costs with an upward trend. Interestingly the results, as will be discussed, can be presented in any document that allows the setting of fonts.

Now what if you wanted to add more display information. What if you wanted to display in the graphic the classification of the store. For example there may be three classifications of stores. One classification is for a facility with a show room for viewing, ordering or returns; a second may provide viewing, ordering or returns plus inventory for purchase; the third may provide all of the previous in addition to servicing. Revenues will vary depending upon the type of store and its size. Unless the viewer filters based on facility type, the display of results can be confusing.

An Approach for Condensed Display of Multi-variable Data

The approach presented uses parameter/variable driven symbolical displays, the use of colour, the cues provided for interpretation, and the ability to quickly convey multi-variable and spatial temporal data. It also is adaptive to changing needs and interpretations and provides ease of display and/or printing on any modern device. It can be used to convey information in a dashboard display to quickly provide information for action. The symbology can be quickly changed and animated to show various trends. Symbols can be chained to convey multiple states of information.

The effort presented is a belated off-shoot of work done in the late 1970's and early 1980's detailed in: CONSIDERATIONS FOR THE DESIGN OF AN INTERACTIVE SPATIAL INFORMATION SYSTEM (ISIS), Moon, G.C; Collins, S.H.; Blair, G, Final report, 15 Jan 1984, Defence and Research Development Canada, DRDKIM No: 84-02706 CANDIS No: 91909. Part of my work during that time was to create an effective means of displaying military symbols and have them updated in real time as new information about military units was obtained. The work also involved using animation to show past movements and likely future movement. Intensity of colour was used to represent reliability of information. So an 'enemy' symbol represented as red would be bright red if the information was certain and a dull red if suspect. Translucency was also a hint for temporal data. There have been other more recent papers on using colour and translucency in a conceptually similar way although not with the same technique described below. (e.g. RESEARCH AND DESIGN OF DYNAMIC SYMBOL IN GIS, Wu Xiaofang, Du Qingyun, Xu Zhiyong, Li Na, School of Resource and Environmental Science, Laboratory of GIS, Ministry of Education, Wuhan University, Wuhan, China, 2005 and EXPLORATORY RESEARCH ON DYNAMIC VISUALIZATION BASED ON SPATIO-TEMPORAL DATABASE, Yang ping, Tang Xinming, Lei Bing, Zhai Liang, Tian Jiaojiao, School of Resource and Environment Science, Wuhan University, Wuhan, China, 2005).

The display device used for ISIS referenced above, even though it was state-of-the-art for the time, was slow and memory was small. A mechanism was needed to efficiently display a large number of symbols to demonstrate a realtime command and control system. Those that are familiar with Military symbols know how dynamic they are. An example using a NATO symbol guide shows:

14  (+) 14 Infantry battalion re-enforced, the X represents infantry (a rounded box would represent armoured). The (+) means re-enforced.

 Armoured infantry unit in a special grouping such as a combat team.

In other words a large number of combinations of basic symbol shapes or numbers and designations. Solving this with the limitations of the day resulted in the use of Hershey fonts, which were early vector fonts (<http://emergent.unpythonic.net/software/hershey>), creating a font that contained the basic shapes of military symbols and then using backspace to graphically overstrike characters in the formation of the symbol. In this way any symbol could be derived on demand while minimizing how many characters were necessary in the base font. In contrast, today there are Unicode representations of pre-built military symbol sets which are not combined; resulting in hundreds of symbols and not all possible permutations.

The use of backspace in today's devices to duplicate the earlier effort is not necessary but the approach

is similar when applied to other applications that could have a large number of different symbols to convey information. That is, create a font containing the foundations of the business symbols as characters and then overstrike the characters to create the appropriate result. All of the methods present for displaying and manipulation fonts are available and inserting a symbol in a text file without cutting and pasting is a matter of changing the font and inserting the characters: 2010/1  p . The previous 'business' symbol illustrates the use of subscript, 11 'base' characters for the symbol and colours. The work discussed below to create this dynamic symbol uses Unicode, an international encoding standard for letters, digits and symbols where an unique number is used for every character.

Combining characters within Unicode is already recognized within the standard to allow the building of complex letters. The combining diacritical marks area (0300-036f) and the combining diacritical marks supplemental area (1DC0-1DFF) are there to provide a mechanism for combining marks to create letters in non-latin languages (although often the combination is available also as a single character representation). Combining marks for symbols (20D0-20FF) provides limited support for dynamic creation by providing a few useful marks such as enclosing keycap, square or circle. Nonetheless, as mentioned, the concept of 'overstrike' is encapsulated in the Unicode standard.

Some have indicated however that it is not adequate for all languages and have provided other means of building sets of user-specific symbols for purposes such as displaying ancient characters or fictitious languages such as Klingon:

Beam me aboard!  . (An Efficient and Dynamic Symbol Code System for the Internationalization of the World Wide Web, Noritaka OSAWA, Graduate School of Information Systems, The University of Electro-Communications, JAPAN, 1998). The approach described does, however, use an applet and a virtual machine. There are advantages of cross-browser and application if a specialized font can be used for a task.

Symbol Representation for Revenue, Cost, Profit and Trend

The Unicode values illustrated in this document are representative for its intended use but are not limited to what is described. There are many instances where the approach can be adapted for use with any set of symbols to be combined. Characters shown are from a created font called DynamicBusinessSymbol (Fontname: DynamicBusinessSymbol Family Name: DynamicBusinessSymbol Weight: Medium) and was created to demonstrate how the approach is used in a business application. The font includes duplicates of base symbols in two ranges. One is the 100 range and the other is the 300 range as shown below. Other ranges can be used however some applications do not fully support the Unicode range of values. The 100 range value is the Latin Extended-A range and should be available for any application stating Unicode values can be used:

Shape	Name	Unicode No
	Container	0320 0120
	No Data	031F 011F
	Percent20-Percent100	0300 – 0304 0100 – 0104
	PercentTop20-PercentTop100	030A – 030E 010A – 010E
	ArrowUp, Arrow45, Arrow90	0305 – 0307 0105 – 0107
	ArrowMinus45, ArrowDown	0308, 0309 0108, 0109

↑ ↗ →	Arrow_Up, Arrow_45, Arrow_90	030F – 0311 010F – 0111
↘ ↓	Arrow_Minus45, Arrow_Down	0312 – 0313 0112 – 0113
1space, 2space	Alternative range character spacing	031d, 031e 011d, 011e
•	1 star, 2 star, 3 star (hexagon)	0125, 0126, 0127

Year, quarter and descriptor can use any font style or character. Any colour or enhancement available for a font can be used.

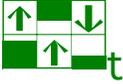


The above uses subscript for 2010/1, that can represent year and quarter; p to indicate profit is a concern; a two star representing in our example a traditional store without in-store servicing; green for two of the measurements that are acceptable and red for profit not meeting expectations. Relief shading can also be used as well to make the symbol stand out more.



Size and / or blinking (if you have a dynamic copy) can be used to bring attention to the symbol. In a text document blinking is likely just distracting but it can be useful in a much more complex display, such as a dashboard or representations on a map of store locations where the blinking quickly brings attention to areas of concern for further examination.

Any colour can be used. Often for monetary display, green is good, red is not good and yellow is a

caution.   2009/3t yy The bars are proportional to numbers being presented to quickly provide an idea of performance. The t could be used to represent target sales. The yy may represent year over year for the previous quarter. The use of t or yy is completely arbitrary and will be the notation appropriate for the language or business practice of the industry.

In a dashboard or on a display with animation, the symbols could be side by side or triggered by a control. Animation could also be used to show the change over time. For example if you are looking at growth you could have a sequence of symbols over a period of time in a line, available under a control or animated over time. Animation can be useful especially on a map display. Projections into the future can also be shown with the same technique. Colour intensity can illustrate the certainty of a

projection. For example, bright is high, dull is suspect  .

The symbols can represent a single location or be used to show regions where the primary colour would be used as a transparent layer to show primary regional indicators and the symbol positioned inside the area to provide more details. Scale dependency can be used so that symbols show more or less information depending on size of the symbol and number of symbols in an area. That is, at a small scale map display where symbols could overlap or details would not be possible to view, display can be set only to show a subset of information such as:  or  or  or  ...

As you can see there are many different possibilities. Clearly the advantage of dynamic symbology is in combining symbolic representations into virtually unlimited numbers of combinations without

having to store each combination as a separate representation. The symbols can also represent two

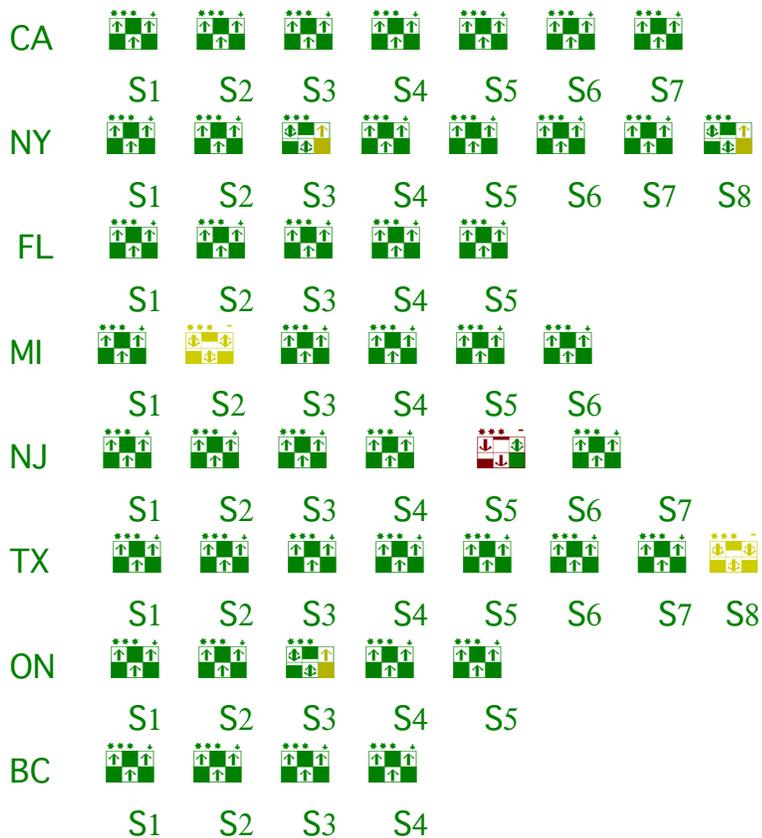
columns:  (character sequence - u0103u0105u0126u0120u010au010fu0120) or one: 

(character sequence – u0104u0105u0125u0120). As a note, some applications that claim unicode support did not support any characters above the 300 range. Also diacritical behaviour appears limited when using multiple characters subsequently, thus any non-diacritical range can be used by having the glyph width control the positioning of the subsequent characters in the business symbol font (which is why the 100 range above works).

A few Examples

Compact Executive Overview

An executive or manager can get a quick assessment of their 50 or more stores:



It is easy to draw attention to any store or other type of outlet as can be seen without requiring very much graphical display space yet displaying more information than is usually summarized for an executive. Drilling into any of the stores is possible.

Informative Interactive Map Displaying

Many possibilities exist when combining the symbols with a map. The following example illustrates the use of symbols like the ones described above in various forms. Some use subscript and superscript, there is a representation of the store type (3 stars), they use colours and size can be significant

depending on use. Also the symbol location can be used to provide additional information as shown with the St. Catherine's balloon window.



Some Details and Samples

A sample of what is presented above can be found at <http://moonlite.office-on-the.net/~moonlite/Moonlite/sample/LiveGeoAddrSymWeb.html>. The sample uses the font created for the symbol. It reads data from an example of a file that contains addresses for facility locations and a representation for the symbol. The sample display uses Google maps, Google geocoding, ELabel JavaScript for displaying the font on a Google map and JavaScript to prepare what is displayed using ELabel. ELabel code can be found at http://code.google.com/p/ludumdare/source/browse/trunk/compo/plugins/rt_simplemap/js/elabel.js?r=82 and a tutorial of its use at <http://econym.org.uk/gmap/elabel.htm> and by looking at the code used in the LiveGeoAddrSym.html sample.

Font Creation and Use

The font DynamicBusinessSymbol was created as a true type font using the open source FontForge program available at <http://fontforge.sourceforge.net/>. For use as a font in text documents the font needs to be locally installed. The web sample above checks if the font is installed locally and will embed the font if it is not available. The free for commercial use Font Squirrel generator at <http://www.fontsquirrel.com/fontface/generator> was used to generate the Microsoft EOT font format. Internet Explorer requires for embedding fonts. Unfortunately, currently stable Internet Explorer

versions either do not respect the glyph spacing set in the font or the generated EOT changes the spacing. I expect it is the former as several attempts were made to have the custom font work within Internet Explorer to no avail including using cufon (<http://cufon.shoqolate.com/generate/>) and sifr (<http://www.sifrgenerator.com/>) that are alternative approaches to embedding.

Internet Explorer 9 (IE9) is becoming much more standards compliant (<http://ie.microsoft.com/testdrive/>) so optimistically I have assumed in browser checks that IE9 will work. In the meantime, Firefox 3.x, Chrome 5.x, Opera 10.x and Realplayer 11.x have successfully displayed the symbols correctly.

Preparing for HTML Display

Styles are used to generate the size and colour of the symbol components. To make them stand out on a map they can be shadowed by using text-shadow modifier or by creating a box for the symbol. Both are used and can be seen in the sample.

The text values that create the symbols are expressed as unicode hexadecimal numbers. The JavaScript code for the sample site has a couple variations. One is using the numbers directly. The sequence `<g class="greenBorderStyle" style="font-size:14pt">\u0103\u0105\u0127\u0120</g>` appears as -  .

Creating the symbol using the unicode numbers directly works well when generated from a program using binning to create the values. An alternative which abstracts the ranges used and allows for creation of symbol examples using English keyboards, provides a lookup to substitute the values. The equivalent green symbol  can be represented as G35rx given the JavaScript for the sample.

How best to prepare for the HTML display depends on what is generating or storing the text used for the symbols.

Where from Here?

Hopefully this presentation will encourage others to use the techniques in their applications and to publish the results. The JavaScript necessary to modify and build upon the effort can be found at the site (<http://moonlite.office-on-the.net/~moonlite/Moonlite/sample/>). The font used is also present and is provide under Creative Commons, Attribution-Share Alike 3.0 Unported - <http://creativecommons.org/licenses/by-sa/3.0/> and is also at the sample site. There is also a commercial offering that uses the font and approach by Moss Solutions, <http://david-moss.com/> . The example at <http://david-moss.com/labs/bizSymbols/storesPerformance.html> shows a spreadsheet to illustrate revenue, cost and profit as well as indicators on performance against expected and previous values of the three. For example the symbol,  shows at a glance a lot of information. The revenue is between 480 to 640 k dollars, down more than 20% from the previous measure, and significantly below target. The cost is between 280 to 420 k dollars, down significantly, and much less than the expected cost. The profit is high and double the previous year.