Nested Venn Diagrams

Nicholas J. Radcliffe
Stochastic Solutions Limited
& Department of Mathematics and Statistics, University of Edinburgh

njr@StochasticSolutions.com

Abstract

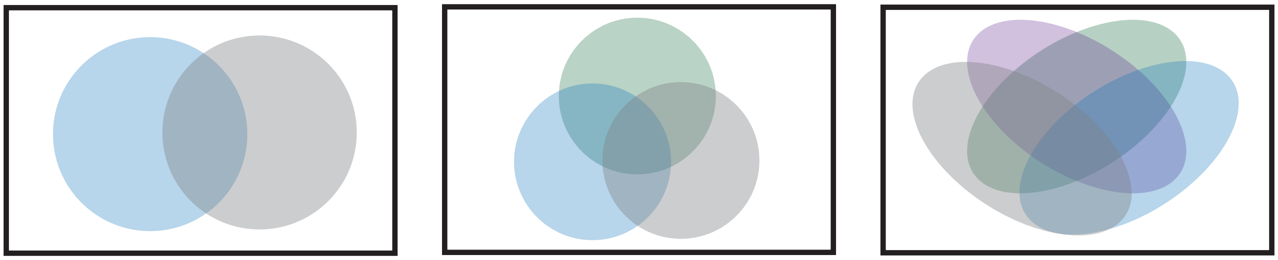
Nesting is used to extend Venn Diagrams from their usual practical limit of three or four sets to a form that readily handles five to eight sets. Although devised independently, such nested diagrams can be viewed as the application to Venn Diagrams of an idea proposed in 1896 by Lewis Carroll, in the context of his alternative set diagrams. Such nesting is rarely if ever used today, but has genuine practical value; ironically, it appears to be more useful in the context of the familar Venn Diagram than when applied to Lewis Carroll's own, less popular, diagrams.

Keywords: Venn Diagams, Lewis Carroll Diagrams, sets, visualization, high-dimensionality, high rank, nesting.

1 Background: Venn Diagrams

Venn Diagrams (Venn, 1880) have been widely used for over a century to show the membership relationships between two and three sets (Figure 1a–b). There are several constructions for Venn Diagrams covering four sets, of which perhaps the most widely used and elegant is shown in Figure 1c. Although Venn constructed a method allowing his diagrams to be extended to an arbitrary number of sets (with his constructions for five and six sets shown in Figure 2), these appear to be used rarely, if ever, in practice.

There are, however, situations in which practical (i.e. comprehensible) diagrams showing the membership relationships between more than four sets would be useful. This paper shows how nesting can be used to extend the practical scope of Venn Diagrams comfortably to six dimensions, and to seven and eight if four-set Venn Diagrams are employed as a basis.



(a) Two sets (b) Three sets (c) Four sets

Figure 1: Venn Diagrams for two, three and four sets

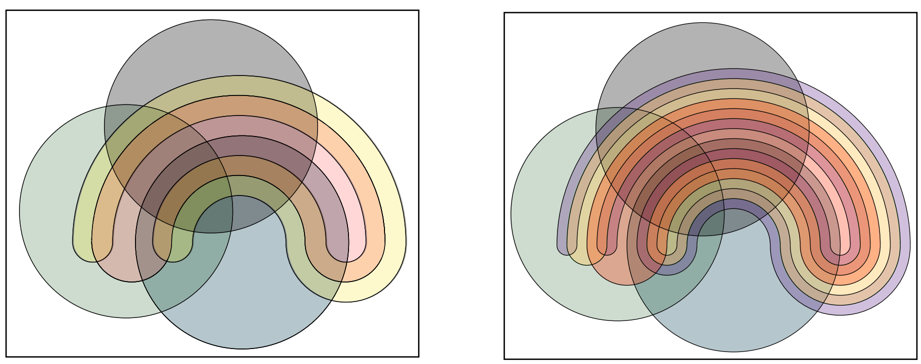


Figure 2: Venn's Constructions for five and six sets. (Adapted from Kopophex, 2008)

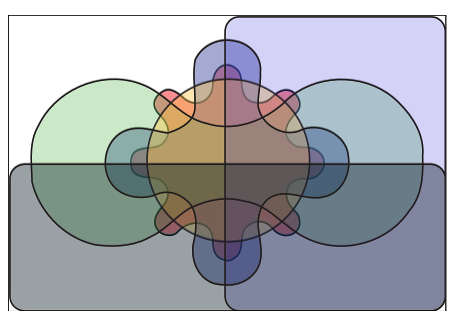
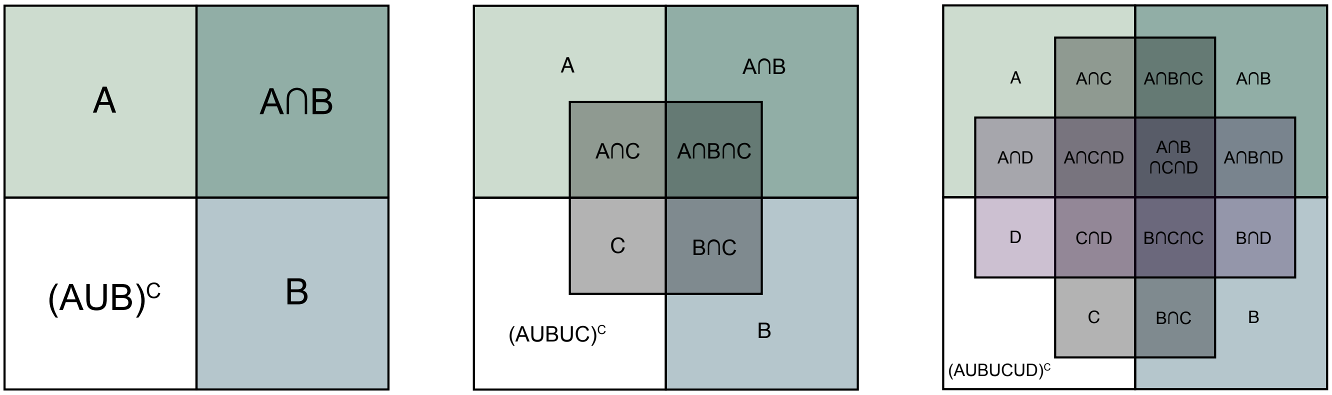


Figure 3: A. W. F. Edwards's 6-set Venn Diagram. (Adapted from Interiot, 2006)



(a) (b) (c)

Figure 4: Lewis Carroll Diagrams for two, three and four sets

A. W. F. Edwards (Edwards, 1989; Edwards, 2004) devised a number of constructions for Venn Diagrams for five, six and seven sets which are widely recognized as easier to use than Venn's constructions for these dimensionalities; the six-set version is illustrated in Figure 3. These, however, still present challenges to comprehension, partly as a result of using different shapes for different sets.

Symmetric Venn Diagrams for larger numbers of sets have also been devised when those numbers are prime; these are surveyed in Ruskey *et al.* (2006).

2 Lewis Carroll Diagrams

An alternative to Venn Diagrams was devised by Lewis Carroll (Carroll, 1896; Edwards, 2004); his constructions for two, three and four sets, are shown in Figure 4. The three- and four-set versions are rarely seen, but the two-set version is widely used in business, most famously in the form of the so-called "Boston Box", or "Growth-Share Matrix" (Figure 5).

It seems likely that the reason Carroll's diagrams for three and four sets have not become more widely used is that people find them harder to read than Venn Diagrams. Although the Carroll diagrams contain strong symmetries, are aesthetically pleasing, and use the same shape for each set, their symmetries actually work against comprehension in some cases; for example, it is not clear from inspection which area corresponds to the intersection of all the sets shown.

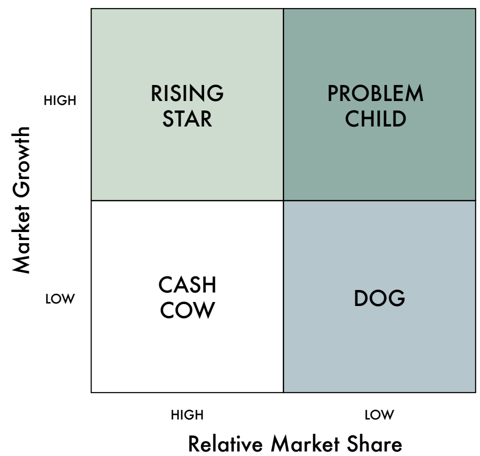


Figure 5: Boston Consulting Group constructed this "matrix" to classify products with respect to market position and market growth. (Here, the two sets would be "High Market Growth" and "Low Market Share".) This matrix is now widely known as *the* "Boston Box", and more generally 2×2 grids (which can be seen to be 2-set Lewis Carroll Diagrams) are often referred to as "Boston Boxes".

It transpires that Carroll also went beyond four sets by using nesting of his diagrams in a manner essentially identical to that proposed for Venn Diagrams in this paper. Figure 6 shows Carroll's eight-set diagram. Although this is again elegant, it is hard to imagine that either the quality or speed of a reader's understanding of the data would be greater using this diagram than using a table. Since the very purpose of diagrams is normally to increase the quality or speed of understanding, this suggests that in practice, nesting Carroll's diagrams is not especially helpful.

3 Nested Venn Diagrams

This paper proposes constructions for Venn-like diagrams for any number of sets from four to eight by single-level nesting of conventional 2-, 3- and 4-set Venn Diagrams. The annotated construction for six sets is shown in Figure 7 and an example of an eight-set Nested Venn Diagram displaying population sizes is shown in Figure 8. It should be clear that for the 6-set case, there is a single, contiguous area for each of the 64 intersections and non-intersections of the six sets, including the complement of the union of the six. This is true of all the Nested Venn Diagrams proposed. Further examples are available from Radcliffe (2010), including sample data. For five sets, there

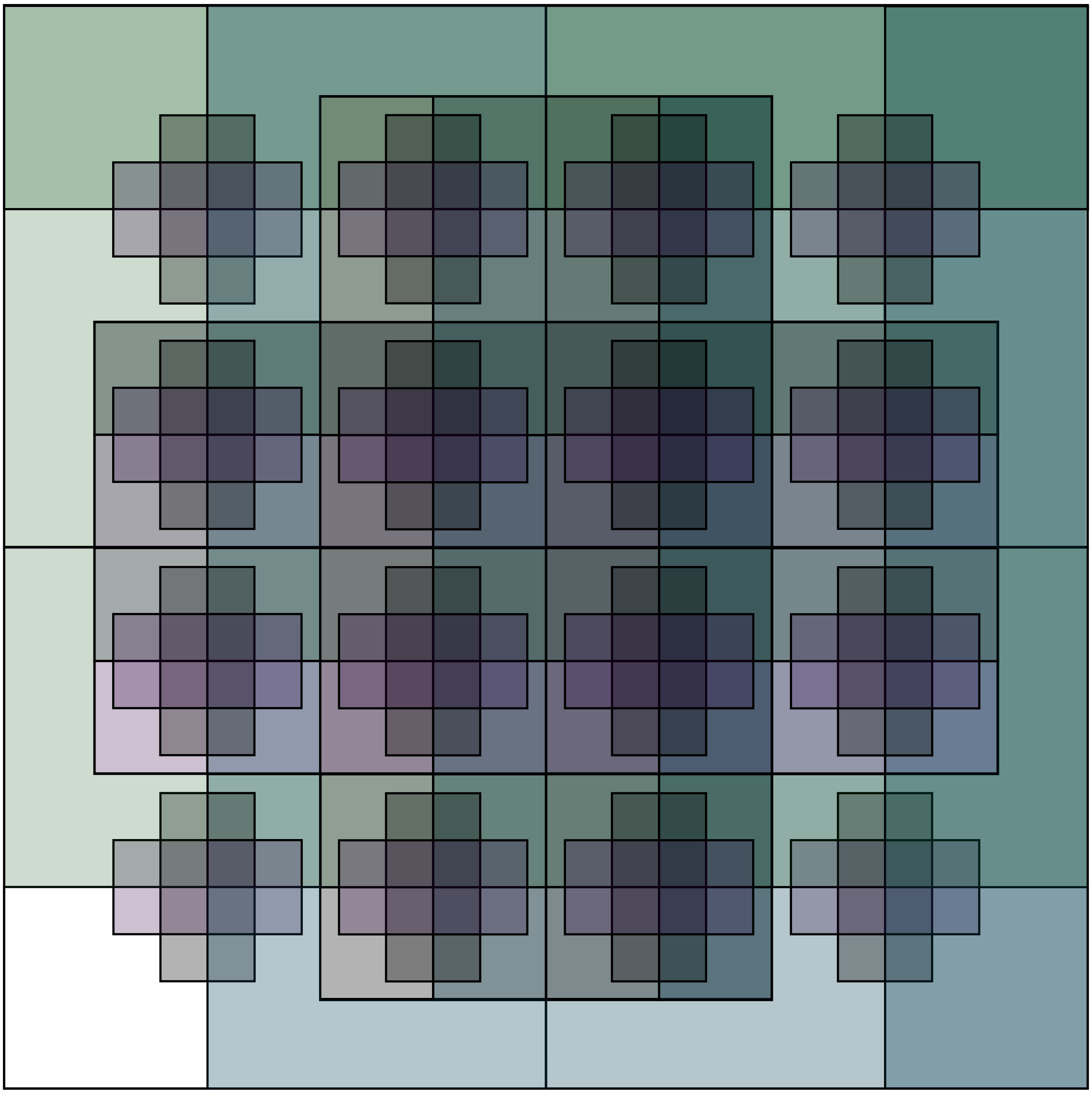


Figure 6: Lewis Carroll Diagram for eight sets.

are obviously reasonable alternatives as to whether to nest small three-set Venn Diagrams in a larger two-set Venn Diagram or small two-set diagrams in a larger three-set diagram. There is a similar choice for seven sets. In practice, we favour using a large three-set diagram in both cases, as illustrated in Figure 9

Obviously, the technique proposed can be trivially extended to nine and more sets either by using more levels of nesting, or by using base Venn Diagrams for more than four sets, but it seems unlikely that such an approach will be useful in practice.

4 Discussion

The Nested Venn Diagrams introduced in this paper were devised when a client innocently asked for a Venn Diagram showing the relationship between visits to six different websites. The author was dimly aware of Venn's constructions for six sets (Figure 2), but after reacquainting himself with them concluded that they would not be suitable as a practical presentation device. Research also revealed Edwards's six-set construction (Figure 3), but while this was clearly more comprehensible than Venn's own six-set diagram, neither the author nor his client felt that it actually aided understanding. Had we at that point discovered Carroll's 6-set diagram, we would certainly have rejected it for the same reason.

In contrast, the diagram actually used (essentially that in Figure 7, but with numbers to show the sizes of sets, as in Figure 8) was of immediate and clear value, allowing patterns to be discerned that were hard to glean from the data in tabular form.

Although the novel *mathematical* content of this paper is limited, we believe that Nested Venn Diagrams have significant *practical* value for displaying the relationships between any number of sets between five and eight. The reactions from people who have seen the diagrams has been largely positive, and people seem to learn to interpret them quickly. As argued above, the purpose of graphical representation is to aid speed of understanding, quality of understanding or both; the best diagrams allow the detection patterns that are not obvious from studying the numbers alone. Edwards (2004) makes a similar point, when he says (on page 47) "In mathematics a good diagram is an invaluable aid to clear reasoning whereas a bad one can seriously mislead." The experience of working with Nested Venn Diagram to date suggests that they meet these tests to a degree that previous diagrams for five to eight sets do not.

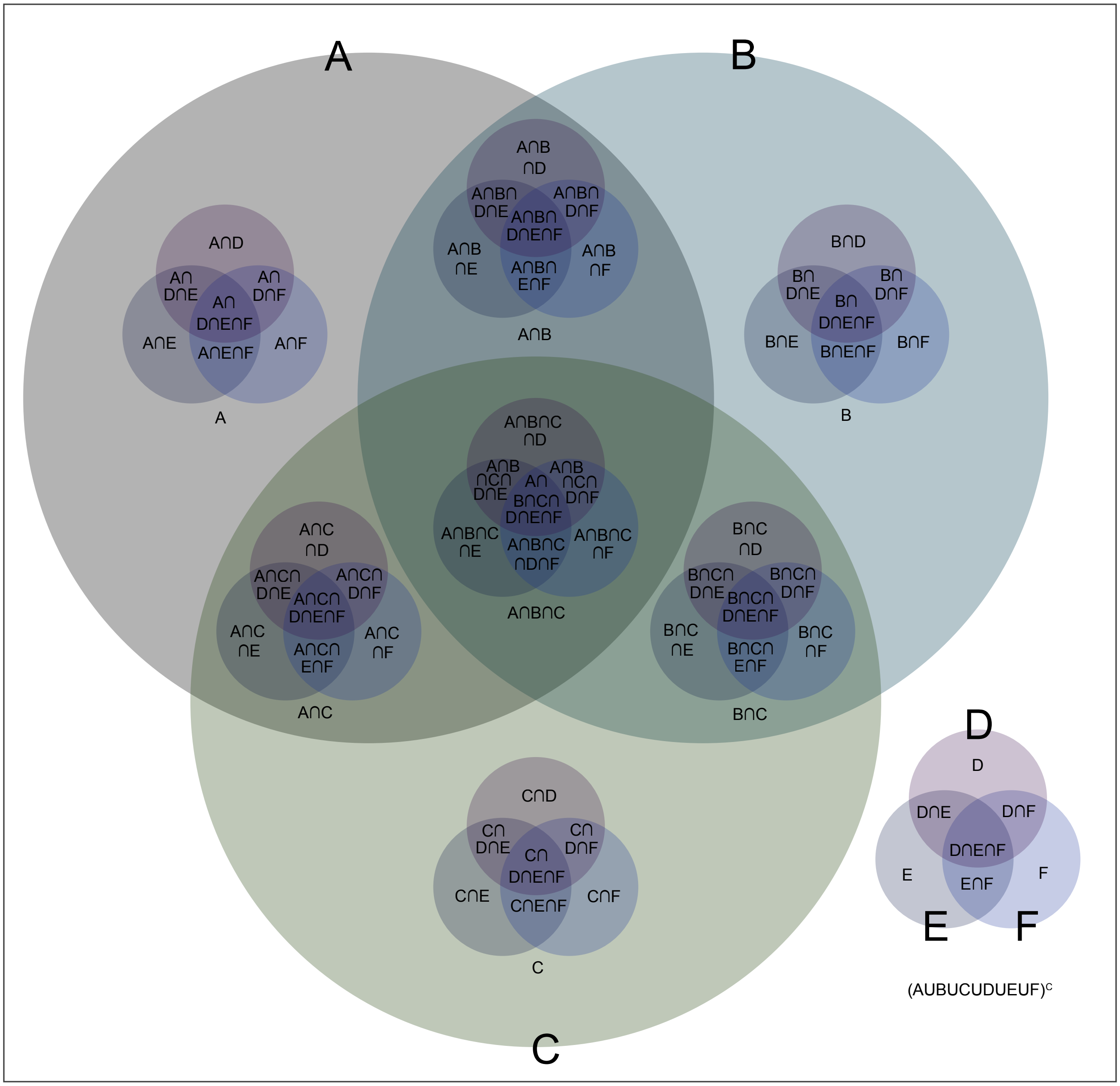


Figure 7: A Nested Venn Diagram for 6 sets.

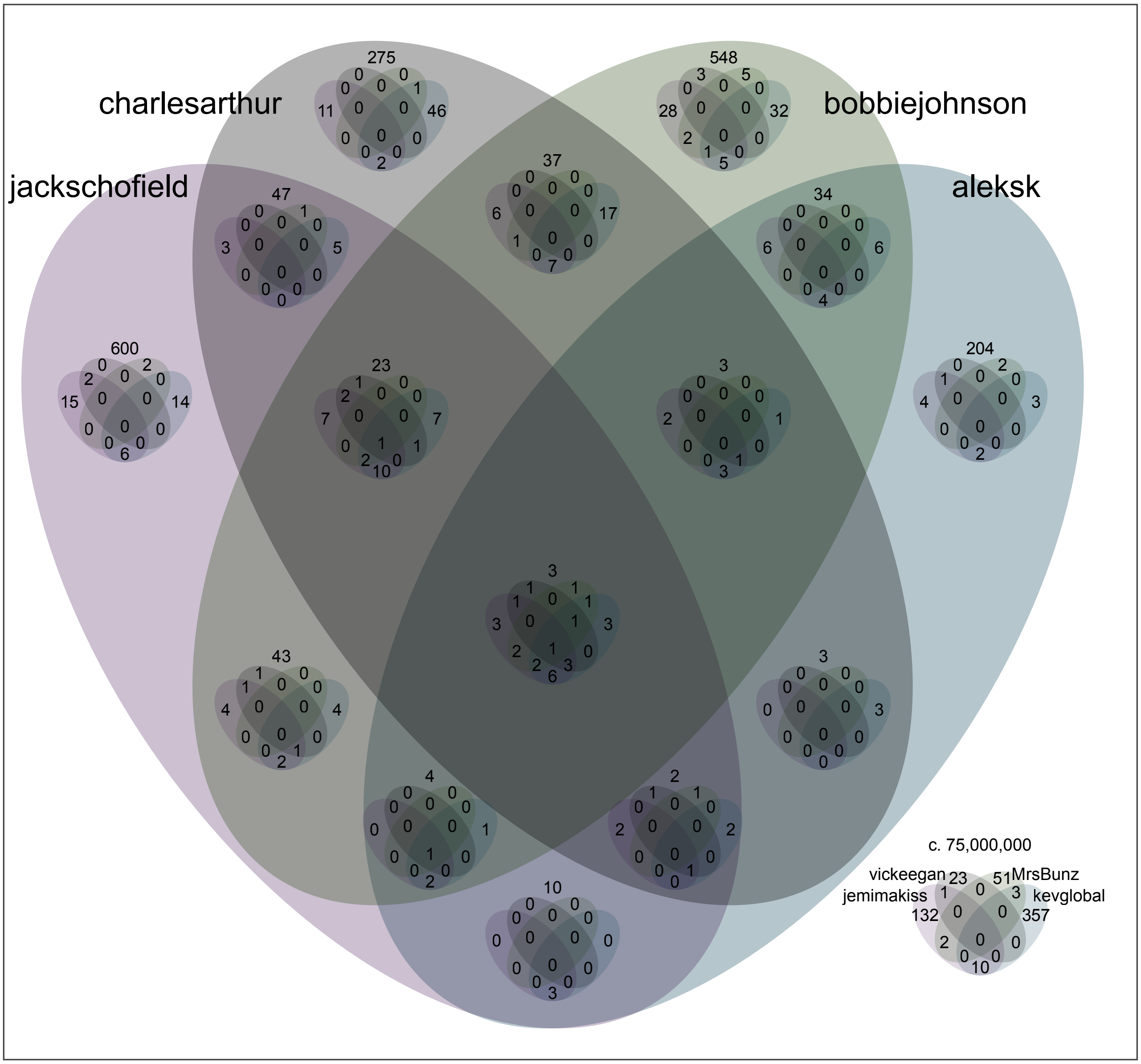


Figure 8: A Nested Venn Diagram for 8 sets. The eight sets correspond to Twitter users as labelled (all members of the Guardian Newspaper's Technology Team). The numbers indicate the number of users they follow (as at 10th February 2010). In the centre, for example, it can be seen that there is a single Twitter user whom all of these eight people follow (this turns out to be The Guardian's Editor, Alan Rusbridger—`@arusbridger`). Similarly, we can see that there are exactly six people followed only by Aleks Krotoski (`@aleksk`) and Jemima Kiss (`@jemimakiss`). In total, however these two users follow 50 users in common, (working in an anticlockwise spiral through the small venn diagrams, starting at the bottom right, $(6+1)+(0+0)+(2+1)+(3+0)+(2+1)+(10+0)+(5+1)+(13+5) = 50$.) The figure of c. 75,000,000, in the universal set but no other, represents the Twitter users not followed by any of the 8 users illustrated, and is based on a January estimate of the total number of Twitter users (Gaudin, 2010).

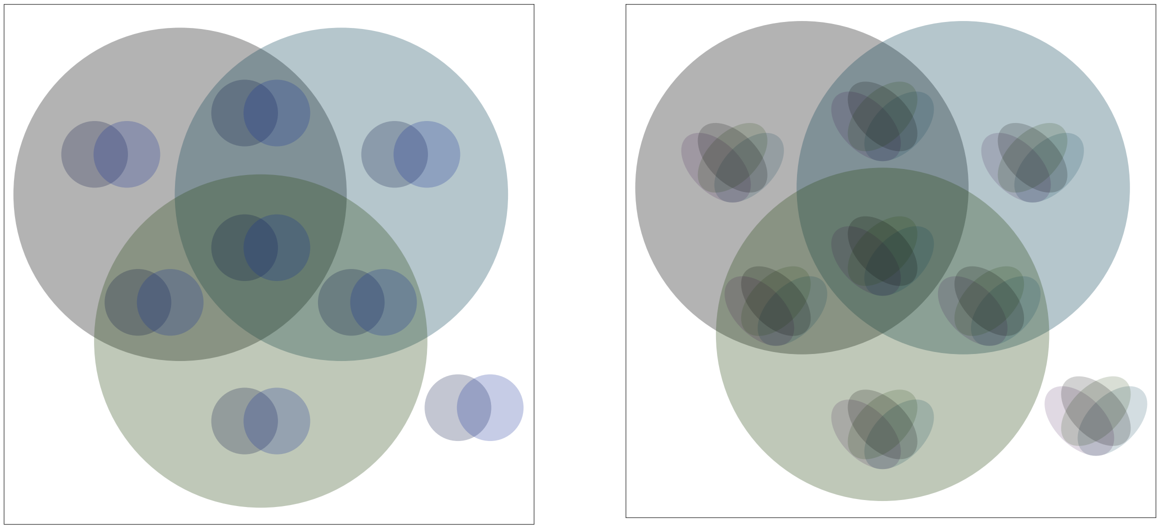


Figure 9: Suggested constructions for five and seven sets.

The Nested Venn Diagram does, of course, suffer some drawbacks compared with conventional Venn Diagrams. The most significant, is that while all the intersections of the large sets are contained within the areas corresponding to those sets, this is not true for the smaller sets (D, E, and F in Figure 7). The various intersections of D, for example, with the other five sets are spread over eight separate areas of the diagram (namely, all of the the "upper" small circles). In this sense, the Nested Venn Diagram construction may be regarded as not a "true" Venn Diagram, and while its interpretation is comparatively straightforward, and to many intuitive, there is clearly some scope for misunderstanding.

As a practical guideline, when the sets have different sizes, it is probably sensible to use the larger circles/ellipses for the larger sets; or more generally, to use the larger circles/ellipes for the sets on which there will be greater focus.

5 Conclusion

The Nested Venn Diagram is proposed as a useful alternative for visualizing and understanding the relationships between a number of sets, particularly when that number is between five and eight.

References

L. Carroll, 1896. *Symbolic Logic*. Macmillan (London).

A. W. F. Edwards, 1989. Venn Diagrams for many sets. *New Scientist*, 121:5156.

A. W. F. Edwards, 2004. *Cogwheels of the mind: the story of Venn diagrams*. The Johns Hopkins University Press.

S. Gaudin, 2010. Twitter now has 75m users; most asleep at the mouse. *Computer World*. January 26, 2010, http://www.computerworld.com/s/article/9148878.

Interiot., 2006. Venn Diagrams. Based on http://en.wikipedia.org/wiki/File:Venn{5,6}.svg Adapted by the author. Originally released under the "Creative Commons Attribution ShareAlike 3.0 license", which is available from http://creativecommons.org/licenses/by-sa/3.0/. Modified version available from http://stochasticsolutions.com/image/Edwards-Venn-six.svg under same license terms.

Kopophex, 2008. Venn Diagrams. http://en.wikipedia.org/wiki/File:Venn{5,6}.svg.

N. J. Radcliffe, 2010. The Nested Venn Diagram. http://scientificmarketer.com/2010/02/nested-venn-diagram.html.

F. Ruskey, C. D. Savage, and S. Wagon, 2006. The search for simple symmetric Venn Diagrams. *Notices of the AMS*, 53(11).

J. Venn, 1880. On the diagrammatic and mechanical representation of proposition and reasoning. *Philosophical Magazine and Journal of Science, Fifth Series*, 9:1–18.