Venn diagrams
Technical details and regression checks

Jonathan Swinton

14th July, 2009

- Try CR for weight=0
- implement not showing dark matter eg Fig 1
- Different choices of first and second sets for AWFE
- Add in the equatorial sets for AWFE
- AWFE-book like figures
- naming of weights for triangles
- likesquares argument for triangles
- central dark matter
- Comment on triangles
- Comment on AWFE return geometry
- text boxes
- use grob objects/printing properly
- proper data handling:
- choose order;
- cope with missing data including missing zero intersection;
- Define weights via names
- graphical parameters
- discuss Chow-Ruskey zero=nonsimple

1 Venn objects

```
> library(Vennerable)
> Vcombo <- Venn(SetNames = c("Female",
+     "Visible Minority", "CS Major"),
+     Weight = c(0, 4148, 409, 604,
+         543, 67, 183, 146))
```

For a running example, we use sets named after months, whose elements are the letters of their names.

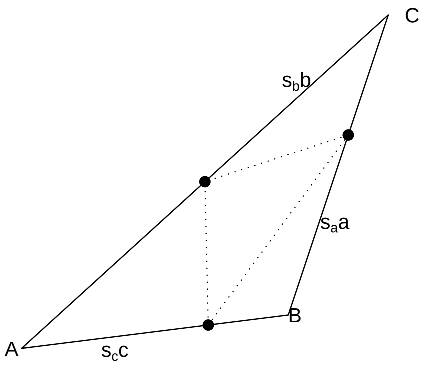
```
> setList <- strsplit(month.name, split = "")
> names(setList) <- month.name
> VN3 <- VennFromSets(setList[1:3])
> V2 <- VN3[, c("January", "February"),
+     ]

> V4 <- VennFromSets(setList[1:4])
> V4f <- V4
> V4f@IndicatorWeight[, ".Weight"] <- 1

> setList <- strsplit(month.name, split = "")
> names(setList) <- month.name
> VN3 <- VennFromSets(setList[1:3])
> V2 <- VN3[, c("January", "February"),
+     ]

> V3.big <- Venn(SetNames = month.name[1:3],
+     Weight = 2^(1:8))
> V2.big <- V3.big[, c(1:2)]

> Vempty <- VennFromSets(setList[c(4,
+     5, 7)])
> Vempty2 <- VennFromSets(setList[c(4,
+     5, 11)])
> Vempty3 <- VennFromSets(setList[c(4,
+     5, 6)])
```



Given a triangle ABC of area Δ and some nonnegative weights $w_a + w_b + w_c < 1$ we want to set s_c, s_a and s_b so that the areas of each of the apical triangles are Δ-proportional to w_a, w_b and w_c. This means

$$\begin{aligned} s_c(1-s_b)bc\sin A &= 2w_a\Delta & (1)\\ s_a(1-s_c)ca\sin B &= 2w_b\Delta & (2)\\ s_b(1-s_a)ab\sin C &= 2w_c\Delta & (3) \end{aligned}$$

So

$$\begin{aligned} s_c(1-s_b) &= w_a & (4)\\ s_a(1-s_c) &= w_b & (5)\\ s_b(1-s_a) &= w_c & (6) \end{aligned}$$

$$\begin{aligned} s_b &= 1-w_a/s_c & (7)\\ s_a &= w_b/(1-s_c) & (8)\\ (s_c-w_a)(1-s_c-w_b) &= s_c(1-s_c)w_c & (9) \end{aligned}$$

$$s_c^2(1-w_c)+s_c(w_b+w_c-w_a-1)+w_a(1-w_b) = 0 \qquad (10)$$

Iff

$$4w_aw_bw_c < (1-(w_a+w_b+w_c))^2 \qquad (11)$$

this has two real solutions between w_a and $1-w_b$.

```
[1] TRUE
```

1.1 Three triangles

2 Chow-Ruskey

See **??**.

2.1 Chow-Ruskey diagrams for 3 sets

The general Chow-Ruskey algorithm can be implemented in principle for an arbitrary number of sets provided the weight of the common intersection is nonzero.

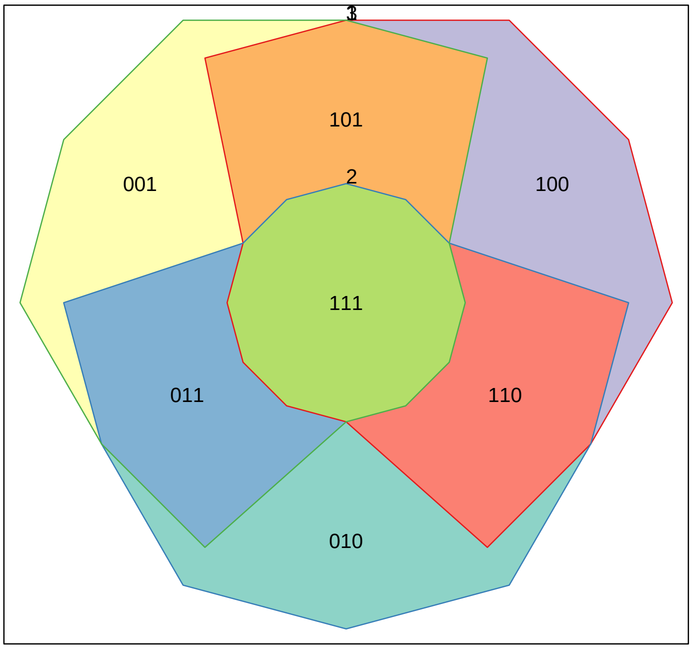
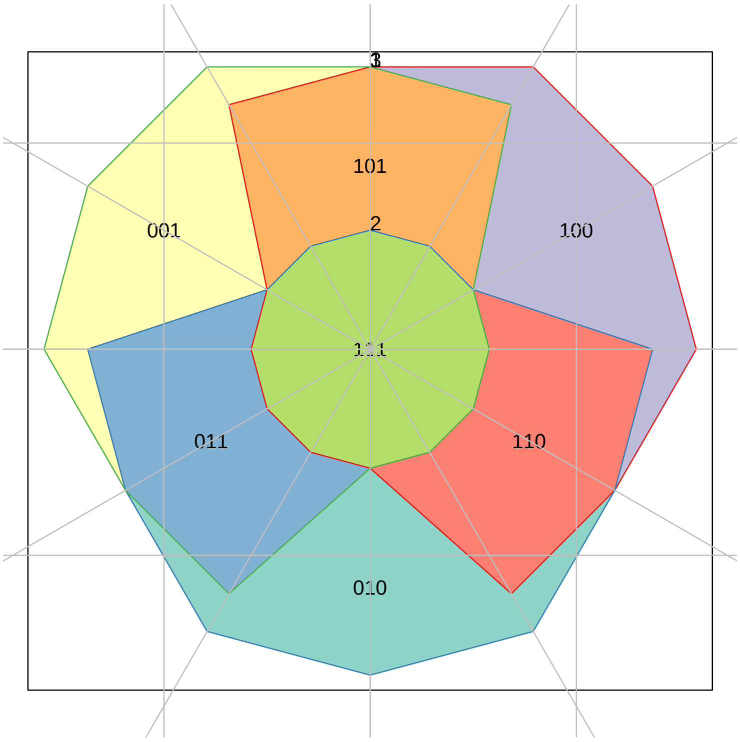


Figure 1: Chow-Ruskey weighted 3-set diagram


3
101
2
001
100
111
011
110
010

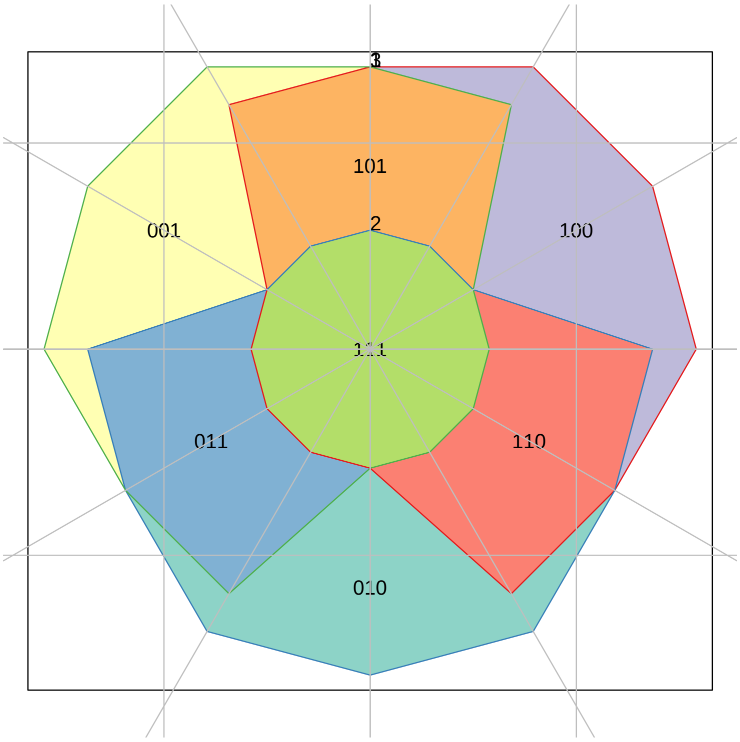


Figure 2: Chow-Ruskey CR3f

2.2 Chow-Ruskey diagrams for 4 sets

Figure 3: Chow-Ruskey weighted 4-set diagram

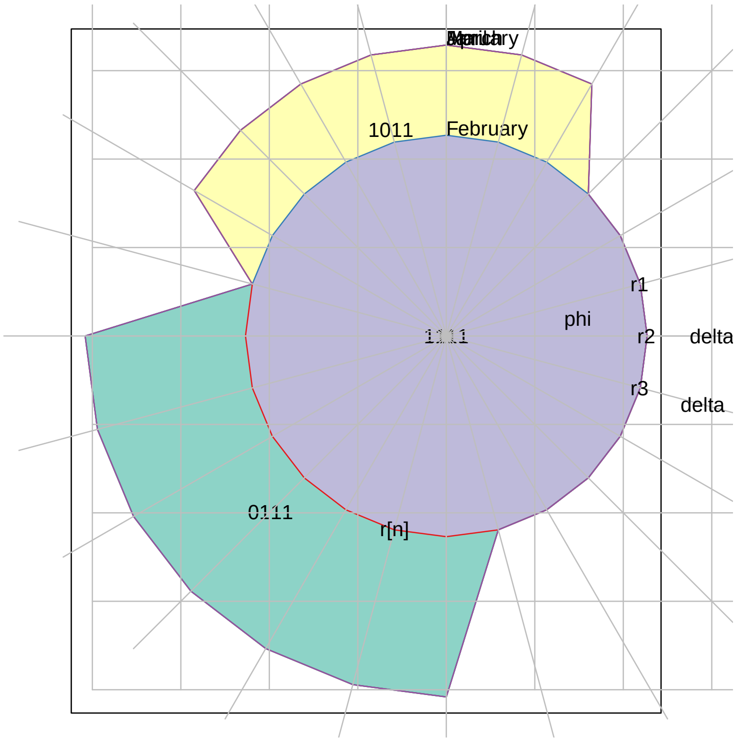


Figure 4: Chow-Ruskey weighted 4-set diagram

The area of the sector $0r_1r_2$ is $\frac{1}{2}r_1r_2\sin\phi$. The area of $0r_1s_2$ is $\frac{1}{2}(r_1(r_2+\delta)\sin\phi)$ and so the area of $r_1r_2s_2$ is $\frac{1}{2}(r_1\delta\sin\phi)$.

The area of $r_2r_2s_2s_3$ is $\frac{1}{2}[(r_3+\delta)(r_2+\delta)-r_3r_2)\sin\phi = \frac{1}{2}[(r_3+r_2)\delta+\delta^2]\sin\phi$.

The total area of the outer shape is

$$
\begin{aligned}
A &= \frac{1}{2}(\sin\phi)\left[(r_1+r_n)\delta+\sum_{k=2}^{n-2}[(r_{k+1}+r_k)\delta+\delta^2]\right] && (12)\\
&= \frac{1}{2}(\sin\phi)\left[(r_1+r_n)\delta+(n-2)\delta^2+\delta\sum_{k=2}^{n-2}[(r_{k+1}+r_k)]\right] && (13)\\
&= \frac{1}{2}(\sin\phi)\left[(r_1+r_2+2r_3+\ldots+2r_{n-2}+r_{n-1}+r_n)\delta+(n-3)\delta^2\right] && (14)
\end{aligned}
$$

so

$$
\begin{aligned}
0 &= c_a\delta^2+c_b\delta+c_c && (15)\\
c_a &= n-3 && (16)\\
c_b &= r_1+r_2+2r_3+\ldots+2r_{n-2}+r_{n-1}+r_n && (17)\\
c_c &= -A/\frac{1}{2}\sin\phi && (18)
\end{aligned}
$$

This is implemented in the compute.delta function.

If all the rs are the same then $c_b = [2(n-3)+4]r = (2n-2)r$.

These constraints are that

$$4w_a w_b w_c < (1-(w_a+w_b+w_c))^2 \tag{19}$$

must hold for both of the sets of numbers

$$\begin{aligned} w_a &= w_{100} \quad (20)\\ w_b &= w_{010} \quad (21)\\ w_c &= w_{001} \quad (22) \end{aligned}$$

and

$$\begin{aligned} w_a &= w_{101}/W \quad (23)\\ w_b &= w_{011}/W \quad (24)\\ w_c &= w_{011}/W \quad (25) \end{aligned}$$

where w_s is the normalised weight of the set with indicator string s and $W = w_{101} + w_{011} + w_{011} + w_{111} = 1-(w_{100}+w_{010}+w_{001})$.

3 This document

Author	Jonathan Swinton
CVS id of this document	Id: Vennville.Rnw 8 2009-07-13 20:50:37Z js229 .
Generated on	14th July, 2009
R version	R version 2.9.0 (2009-04-17)