

Venn Diagrams with `gplots`

Steffen Möller

December 14, 2012

Venn diagrams¹ allow for a quick overview on the number of elements that multiple sets share. And when those elements are representing traits of real objects, like observations in biomedical sciences/marketing/..., then this may direct researchers to further investigations/decisions.

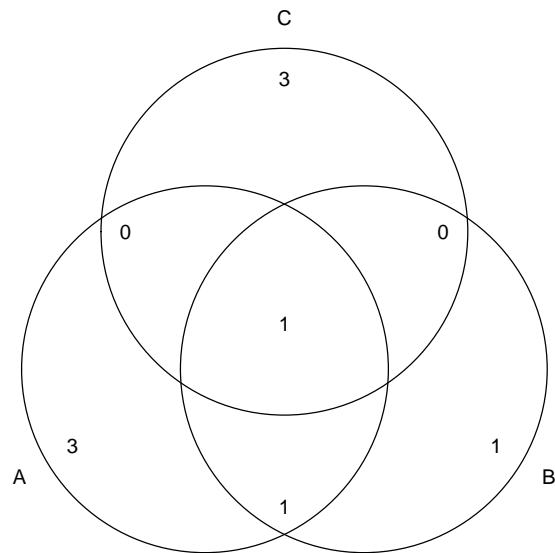
The `gplots` package provides Venn diagrams for up to five sets. The R code to produce the diagrams is not complicated. The plot function behaves always the same, depending only on the number of overlapping circles to draw. Its input is a table that is produced by another function. The function `venn()` calls one after the other and is the only one to be seen by the user. The values shown are returned invisibly.

The `venn()` function accepts either a list of sets as an argument, or it takes a binary matrix, one column per set, indicating for every element, one per row, the membership with every set.

The common form with overlapping circles only works with up to three sets, as seen here:

```
> library(gplots)
> venn( list(A=1:5,B=4:6,C=c(4,8:10)) )
```

¹http://en.wikipedia.org/wiki/Venn_diagram



The names of columns or the list elements are the set names. To squeeze extra circles in, those circles need to become ellipses. This works for four sets

```
> v.table<-venn( list(A=1:5,B=4:6,C=c(4,8:10),D=c(4:12)) )
> print(v.table)
```

```
      num A B C D
0000    0 0 0 0 0
0001    3 0 0 0 1
0010    0 0 0 1 0
0011    3 0 0 1 1
0100    0 0 1 0 0
0101    1 0 1 0 1
0110    0 0 1 1 0
0111    0 0 1 1 1
1000    3 1 0 0 0
1001    0 1 0 0 1
1010    0 1 0 1 0
1011    0 1 0 1 1
1100    0 1 1 0 0
1101    1 1 1 0 1
1110    0 1 1 1 0
1111    1 1 1 1 1
attr(,"class")
[1] "venn"
```

and maybe even more impressively also for five.

The man page of *venn()* lists options to change the appearance of the plots,

e.g., the names of the sets may be omitted and sizes changed. However, there is ample of opportunity to extend the functionality of this package. To mind come

- more dimensions (next)
- colors
- variation of size of circles with the number of members the set has
- density plot rather than numbers, identification of individual entries

The prime personal interest is more in the increase of dimensions. Please send patches for features you are most interested in.