

```

1  --- 768
2  --- 384
3  --- 192
4  --- 96
5  --- 48 (positive)
6  --- 48 (positive)
7  --- 96
8  --- 48 (positive)
9  --- 48 (negative)
10 --- 192
11 --- 96
12 --- 48 (positive)
13 --- 48 (negative)
14 --- 96
15 --- 48 (positive)
16 --- 48 (positive)
17 --- 384
18 --- 192
19 --- 96
20 --- 48 (negative)
21 --- 48 (negative)
22 --- 96
23 --- 48 (negative)
24 --- 48 (negative)
25 --- 192
26 --- 96
27 --- 48 (negative)
28 --- 48 (negative)
29 --- 96
30 --- 48 (negative)
31 --- 48 (negative)
32
33 local R = require
34 local the, eggs, lib = R"the", R"egs", R"lib"
35 local per, cos, norm, o, fmt, rnds = lib.per, lib.cosine, lib.norm, lib.o, lib.fmt, lib.rnds
36 local map, any, many, sort, upl = lib.map, lib.any, lib.many, lib.sort, lib.upl
37
38 local cluster={}
39 function cluster.new(top, egsl, i, lefts, rights)
40   egsl = egsl or top
41   i = {egs=egsl, top=top, rank=0}
42   lefts, rights, i.left, i.right, i.border, i.c = cluster.half(top, egsl.rows)
43   if #egsl.rows >= 2*(#top.rows)^the.leaves then
44     if #lefts.rows < #egsl.rows then
45       i.lefts = cluster.new(top, lefts)
46       i.rights = cluster.new(top, rights) end end
47   return i end
48
49 --- 5 11 10 11
50
51 function cluster.show(i, pre, front)
52   pre = pre or ""
53   local front = fmt("%s%s", pre, #i.egs.rows)
54   if cluster.leaf(i)
55   then print(fmt("%-20s", front, o(rnds(egs.mid(i.egs, i.egs.cols.y))))))
56   else print(front)
57   if i.lefts then cluster.show(i.lefts, " | ".pre)
58   if i.rights then cluster.show(i.rights, " | ".pre) end end end end
59
60 function cluster.leaf(i) return not (i.lefts or i.rights) end
61
62 --- 11 10 11
63
64 function cluster.dist(eg1, row1, row2)
65   local function sym(c, x, y) return x==y and 0 or 1 end
66   local function num(c, x, y)
67     if x=="?" then y = norm(c.lo, c.hi, y); x=y<.5 and 1 or 0
68     elseif y=="?" then x = norm(c.lo, c.hi, x); y=x<.5 and 1 or 0
69     else x, y = norm(c.lo, c.hi, x), norm(c.lo, c.hi, y) end
70     return math.abs(x-y) end
71   local function dist(c, x, y)
72     return x=="?" and y=="?" and 1 or (c.nump and num or sym)(c, x, y) end
73   local d, n = 0, #eg1.cols.x
74   for key, c in pairs(eg1.cols.x) do d=d+dist(c, row1[c.at], row2[c.at])^the.p end
75   d return (d/n)^(1/the.p) end
76
77 function cluster.neighbors(eg1, r1, rows)
78   return sort(map(rows or eg1.rows,
79     function(r2) return {cluster.dist(eg1, r1, r2), r2} end), upl) end
80
81 --- 5 10 11 10 11
82
83 function cluster.half(eg1, rows)
84   local project, far, some, left, right, c, lefts, rights, border
85   rows = rows or eg1.rows
86   far = function(r, t) return per(cluster.neighbors(eg1, r, t), the.far)[2] end
87   project = function(r)
88     return {cos(cluster.dist(eg1, left, r),
89       cluster.dist(eg1, right, r),
90       c),
91     r} end
92   some = many(rows, the.some)
93   left = far(any(some), some)
94   right = far(left, some)
95   c = cluster.dist(eg1, left, right)
96   lefts, rights = egs.clone(eg1), egs.clone(eg1)
97   for n, projection in pairs(sort(map(rows, project), upl)) do
98     if n==#rows/2 then border = projection[1] end
99     egs.add(n <= #rows/2 and lefts or rights, projection[2]) end
100   return lefts, rights, left, right, border, c end
101
102 return cluster

```