

```

1 local your = {} -- user settings (may be changes from command-line)
2 local our = { -- system settings (controlled internal to code)
3     help = {}
4 }
5 lua keys5.lua [OPTIONS]
6 (c)2022, Tim Menzies <tim@ieee.org>, unlicense.org
7
8 --Debug true
9 --far .9
10 --file ../data/auto93.csv
11 --h false
12 --p 2
13 --round 2
14 --seed 10019
15 --Some 512
16 --todo all }}}
17
18 our.defaults={} -- defaults, built from "our.help", written to "your"
19 our.fail=0 -- failure counts (returned on exit)
20 our.oid=0 -- unique id for objects
21 our.go={} -- list of enabled tests
22 out.nogo={} -- list of disabled test
23
24 our.b4={} -- globals known, pre-code. used to find stray globals
25 for k,_ in pairs(_ENV) do our.b4[k]=k end
26
27 local add, any, asserts,coerce, col, copy, csv, defaults, dist
28 local fmt, klass, map, main, new,o, push, rand, randi, rnd, rnds
29 local same, slots, sort
30
31 function klass(s, it)
32     it = {is=s, toString=o}
33     it._index = it
34     return setmetatable(it, {_call=function(_,...) return it.new(...) end}) end
35
36 local EG,EGS,NUM,SYM = klass"EG", klass"EGS", klass"NUM", klass"SYM"
37
38 -----
39 function NUM.new(n,s)
40     return col(n,s, new(NUM,{mu=0, m2=0, lo=math.huge, hi=-math.huge})) end
41
42 function NUM.add(i,x, d)
43     if x=="?" then
44         i.n = i.n+1
45         d = x - i.mu
46         i.mu = i.mu + d/i.n
47         i.m2 = i.m2 + d*(x-i.mu)
48         i.lo = math.min(i.lo,x); i.hi = math.max(i.hi,x) end
49     return x end
50
51 function NUM.dist(i,x,y)
52     if x=="?" then y = i:norm(y); x=y>.5 and 0 or 1
53     elseif y=="?" then x = i:norm(x); y=x>.5 and 0 or 1
54     else x,y = i:norm(x), i:norm(y) end
55     return math.abs(x-y) end
56
57 function NUM.div(i) return i.n<2 and 0 or (i.m2/(i.n-1))^0.5 end
58
59 function NUM.mid(i) return i.mu end
60
61 function NUM.norm(i,x) return i.hi-i.lo<1E-9 and 0 or (x-i.lo)/(i.hi-i.lo) end
62
63 -----
64 function SYM.new(n,s) return col(n,s, new(SYM,{has={},most=0,mode=nil})) end
65
66 function SYM.add(i,x,count)
67     count = count or 1
68     i.has[x] = count + (i.has[x] or 0)
69     if i.has[x] > i.most then i.most,i.mode = i.has[x], x end end
70
71 function SYM.dist(i,x,y) return x==y and 0 or 1 end
72
73 function SYM.div(i, e)
74     e=0; for _,n in pairs(i.has) do e=e-n/i.n*math.log(n/i.n,2) end; return e end
75
76 function SYM.mid(i) return i.mode end
77
78 -----
79 function EG.new(t) return new(EG, {cooked={}, has=t}) end
80
81 function EG.better(eg1,eg2,egs)
82     local s1,s2,e,n,a,b = 0,0,10,#egs.cols.y
83     for _,c in pairs(egs.cols.y) do
84         a = c:norm(eg1.has[c.at])
85         b = c:norm(eg2.has[c.at])
86         s1 = s1 - e^(c.w * (a-b)/n)
87         s2 = s2 - e^(c.w * (b-a)/n) end
88     return s1/n < s2/n end
89
90 function EG.cols(i,cols) return map(cols,function(x) return i.has[x.at] end) end
91
92 function EG.dist(i,j,egs, d)
93     d = 0
94     for _,c in pairs(egs.cols.x) do
95         d = d + dist(i.has[c.at], j.has[c.at], c)^your.p end
96     return (d/(1E-31 + #egs.cols.x))^(1/your.p) end
97
98
99
100 function EGS.new(i) return new(EGS, {rows={}, cols={all={},x={},y={}}}) end
101
102 function EGS.add(i,eg, now,where)
103     eg = eg.has and eg.has or eg -- If data is buried inside, the expose it.
104     if #i.cols.all>0 then
105         push(i.rows, EG(map(i.cols.all, function(c) return add(eg[c.at],c) end)))
106     else
107         for at,s in pairs(eg) do -- First row. Create the right columns
108             now = col(at,s, (s:find"^[A-Z]" and NUM or SYM)())
109             push(i.cols.all, now)
110             if not s:find"." then
111                 where = (s:find"-" or s:find"+") and i.cols.y or i.cols.x
112                 push(where, now) end end end
113
114 function EGS.clone(i,init, j)
115     j = EGS()
116     j:add(map(i.cols.all, function(col) return col.txt end))
117     for _,x in pairs(init or {}) do j:add(x) end
118     return j end
119
120 function EGS.from(t, i)
121     i=i or EGS(); for _,eg in pairs(t) do i:add(eg) end; return i end
122
123 function EGS.read(c, i)
124     i=i or EGS(); for eg in csv(c) do i:add(eg) end; return i end
125
126 function EGS.far(i,egl, fun,tmp)
127     fun = function(eg2) return {eg2, egl:dist(eg2,i)} end
128     tmp = #i.rows > your.Some and any(i.egs, your.Some) or i.egs
129     tmp = sort(map(tmp, fun), function(a,b) return a[2] < b[2] end)
130     return table.unpack(tmp[#tmp*your.far//1] ) end
131
132 function EGS.branch(i)
133     local zero,one,two,ones,twos,both,a,b,c
134     zero = any(i.egs)
135     one = i:far(zero)
136     two,c = i:far(one)
137     ones,twos,both = {}, {}, {}
138     for _,eg in pairs(i.egs) do
139         a = eg:dist(one,i)
140         b = eg:dist(two,i)
141         push(both, {(a^2 + c^2 - b^2) / (2*c),eg}) end
142     for n,pair in pairs(sort(both, function(a,b) return a[1] < b[1] end)) do
143         push(n <= #both//2 and ones or twos, pair[2]) end
144     return ones, twos end

```

```

142
143
144 function add(x,i) if x=="?" then i.n = i.n+1; i:add(x) end; return x end
145
146 function any(t, n)
147   if not n then return t[randi(1,#t)] end
148   u={};for j=1,n do push(u,any(t)) end; return u end
149
150 function asserts(test,msg)
151   msg=msg or ""
152   if test then return print(" PASS:".msg) end
153   our.fails = our.fails+1
154   print(" FAIL:".msg)
155   if your.Debug then assert(test,msg) end end
156
157 function coerce(x)
158   if x=="true" then return true elseif x=="false" then return false end
159   return tonumber(x) or x end
160
161 function col(at,s,i)
162   i.n, i.at, i.txt = 0, at or 0, s or ""
163   i.w = i.txt:find("-" and -1 or 1)
164   return i end
165
166 function copy(t,u)
167   u={}; for k,v in pairs(t) do u[k]=v end
168   return setmetatable(u, getmetatable(t)) end
169
170 function csv(file, x,row)
171   function row(x, t)
172     for y in x:gsub("%s+", ""):gmatch("[^,]+" do push(t,coerce(y)) end
173     return t
174   end
175   file = io.input(file)
176   return function()
177     x=io.read(); if x then return row(x, {}) else io.close(file) end end end
178
179 function defaults(help_string, t,fun)
180   t = {}
181   function fun(flag,x)
182     for n,txt in ipairs(arg) do
183       if flag:match("^"..txt:sub(2).."*")
184       then x = x=="false" and "true" or x=="true" and "false" or arg[n+1] end end
185       t[flag] = coerce(x)
186     end
187     help_string:gsub("(^%s+)[^%n]*%s([%s+])", fun)
188     return t end
189
190 function dist(x,y,i) return x=="?" and y=="?" and 1 or i:dist(x,y) end
191
192 function fmt(...) return string.format(...) end
193
194 function map(t,f, u)
195   u = {};for k,v in pairs(t) do push(u, (f or same)(v)) end; return u end
196
197 function new(mt,x)
198   our.oid = our.oid+1; x._oid = our.oid -- Everyone gets a unique id.
199   return setmetatable(x,mt) end -- Methods now delegate to 'mt'.
200
201 function o(t)
202   local u,key
203   key= function(k) return fmt(":%s %s", k, o(t[k])) end
204   if type(t) ~= "table" then return tostring(t) end
205   u = #t>0 and map(t,o) or map(slots(t),key)
206   return (t._is or "").."["..table.concat(u, " ").".."]" end
207
208 function main()
209   our.defaults = defaults(our.help)
210   our.fails = 0
211   your = copy(our.defaults)
212   if your.h then os.exit(print(our.help)) end
213   for _,one in pairs(your.todo=="all" and slots(our.go) or {your.todo}) do
214     your = copy(our.defaults)
215     our.go[one]()
216   end
217   for k,v in pairs(_ENV) do
218     if not our.b4[k] then print("/rogues",k,type(v)) end end
219   os.exit(our.fails) end
220
221 function push(t,x) table.insert(t,x); return x end
222
223 function rand(lo,hi)
224   your.seed = (16807 * your.seed) % 2147483647
225   return (lo or 0) + ((hi or 1) - (lo or 0)) * your.seed / 2147483647 end
226
227 function randi(lo,hi) return math.floor(0.5 + rand(lo,hi)) end
228
229 function rnd(x,d, n)
230   if type(x)~="number" then return x end
231   n=10^(d or your.round)
232   return math.floor(x*n+0.5)/n end
233
234 function rnds(t,d) return map(t,function(x) return rnd(x,d) end) end
235
236 function same(x,...) return x end
237
238 function slots(t, u)
239   u={ }
240   for k,_ in pairs(t) do if tostring(k):sub(1,1) ~= "_" then push(u,k) end end
241   return sort(u) end
242
243 function sort(t,f) table.sort(t,f); return t end

```

```

244
245 local go=our.go
246 function go.num( m,n)
247   m=NUM()
248   for i=1,10 do add(i,m) end
249   n = copy(m)
250   for i=1,10 do add(i,n) end
251   asserts(2.95 == rnd(n:div()), "sd ok") end
252
253 function go.egs( egs)
254   egs = EGS.read(your.file)
255   asserts(egs.cols.y[1].hi==5140, "most seen") end
256
257 function go.clone( egs1,egs2,s1,s2)
258   egs1 = EGS.read(your.file)
259   s1 = o(egs1.cols.y)
260   egs2 = egs1:clone(egs1.rows)
261   s2 = o(egs2.cols.y)
262   asserts(s1==s2, "cloning works") end
263
264 function go.dist()
265   local egs,egl,dist,tmp,j1,j2,d1,d2,d3,one
266   egs = EGS.read(your.file)
267   egl = egs.rows[1]
268   dist = function(eg2) return {eg2,egl:dist(eg2,egs)} end
269   tmp = sort(map(egs.rows, dist), function(a,b) return a[2] < b[2] end)
270   one = tmp[1][1]
271   for j=1,10 do
272     j1 = randi(1,#tmp)
273     j2 = randi(1,#tmp)
274     if j1>j2 then j1,j2=j2,j1 end
275     d1 = tmp[j1][1]:dist(one,egs)
276     d2 = tmp[j2][1]:dist(one,egs)
277     asserts(d1 <= d2, "distance") end end
278
279 main()

```