```
csc710sbse: HW5:Theisen
   Oct 05, 14 21:41
                                                                                                                                                     Page 1/1
#Structure from SA Lecture
import sys,re,random,math
sys.dont_write_bytecode = True
from options import *
from utils import *
from sk import *
myOpt = Options()
class Analyzer:
  n = 50
old = [1 for i in range (0, n)]
new = [1 for i in range (0, n)]
   era_lives = myOpt.era_lives;
  def bettered(self, new, old):
     def quartiles(value):
        return value*.25, value*.5, value*.75
      def betterifless():
        p1, median1, p3 = quartiles(new)
IQR1=p3-p1
        p1, median2, p3 = quartiles(old)
IQR2=p3-p1
return median1<median2, IQR1<IQR2
      def same(): return al2(new, old)≤0.56
      betterMedian, betterIQR = betterifless()
      return betterMedian, betterIQR, same()
   def EraStop(self, lst):
     self.old = self.new
self.new = lst
      out = False
     betterMedian = False
betterIOR = False
      same = False
      #print self.old
#print self.new
      oldQ1, oldMedian, oldQ3 = quartiles(self.old)
newQ1, newMedian, newQ3 = quartiles(self.new)
if newMedian < oldMedian:
     betterMedian = True
if new03 - new01 < old03 - old01:
betterIQR = True
if al2(self.new, self.old) ≤ myOpt.al2_test:
         same = True
     if (same \( \backsim \) betterIQR) \( \backsim \) same \( \backsim \) betterMedian):
    out = False
      #bettered
     elif (¬ same ∧ betterMedian):
  out = True
     if out:
    self.era_lives += 1
     self.era_lives -= 1
if self.era_lives = 0:
    #print "Early Era Termination!"
    return True
      else:
        return False
#from menzies code
def median(lst,ordered=False):
  if ¬ ordered: lst= sorted(lst)
  n = len(lst)
  p = n//2
if n % 2: return lst[p]
  q = p - 1
q = max(0,min(q,n))
return (lst[p] + lst[q])/2
def quartiles(lst):
    q1 = lst[int(len(lst)*.25)]
    med = median(lst, False)
  q3 = lst[int(len(lst)*.5)]
return q1, med, q3
```

```
#From Class Discussion 8/26/2014
from _future__ import division
import pys,re,random,math
import numpy as np
    sys.dont_write_bytecode = True

class Options:
    #Clobals
    debug = False
    seed = 1
    era_lives = 3
    al2_test = 0.6

#MaxWalkSat options
    mws_prob = 0.75
    mws_maxTries = 500
    mws_maxChanges = 500
    mws_threshold = .000001
    mws_elices = 10

#Simulated Annealing options
    sa_kmax = 500
    sa_cooling = .6

#Genetic Algorithm options
    ga_gen_list = [10]
    ga_pop_size = 50
    ga_crossover = .6

def printClobals(self):
    print "Red:", self.seed, "Lives:", self.era_lives
    print "Al2test:", self.al2_test
```

Page 1/1

```
csc710sbse: HW5:Theisen
     Sep 23, 14 13:25
                                                                                                                                                                                                            Page 1/1
from __future__ import division
import sys,re,random,math
import numpy as np
sys.dont_write_bytecode = True
 from sk import *
def rdiv8():
  rdivDemo([
    ["novels", 287, 332, 443, 711, 534],
    ["kids", 23, 18, 16, 20, 21],
    ["magazine", 112, 98, 43, 63, 82],
    ["cookbooks", 232, 180, 32, 53, 78],
    ])
  rdiv8()
```

Oct 05, 14 21:51	csc710sbse: HW5:Theisen	Page 1/4
Model Name: Fonseca Searcher Name: SA Seed: 1 Lives: 3 Al2test: 0.6 SA Options: KMAX: 500 Cooling: 0.6 Time to run (s): 2.916281 Runs: 5		
Average per run (s): 0.5832562 (*), 0.83001,		
Model Name: Fonseca Searcher Name: MWS Seed: 1 Lives: 3 Al2test: 0.6 MaxWalkSat Options: Prob: 0.75 MaxTries: 500 MaxChanges 500 Threshold: 1e-06 Slices: 10 Time to run (s): 4.12327		
No valid runs!		
Average per run (s): 3.218167	0.51859, 0.53717, 0.68139, 0.93113	
rank , name , med , iqr 1 , GA , 54 , 16 (- 1 , SA , 94 , 9 (*), 0.50, 0.52, 0	1.54, 0.68, 0.93
Model Name: Schaffer Searcher Name: SA Seed: 1 Lives: 3 Al2test: 0.6 SA Options: KMAX: 500 Cooling: 0.6 Time to run (s): 0.138624 Runs: 5 Average per run (s): 0.0277248 (************************************	0.00000, 0.00000, 0.00000, 0.00003	
Average per run (s): 0.151333		
	0.00000, 0.00000, 0.00003, 0.00006	

rank , name , med , iqr	csc710sbse: HW5:Theisen		Page 2/4
			-
1 , SA , 0 , 0 (1 , MWS , 0 , 0 (1 , GA , 0 , 0 (*	0.00, 0.00, 0.00, 0.00, 0.00, 0.00,	0.00 0.00 0.00
Model Name: Kursawe			
Searcher Name: SA Seed: 1 Lives: 3 Al2test: 0.6 SA Options:			
KMAX: 500 Cooling: 0.6 Time to run (s): 2.531 Runs: 5			
Average per run (s): 0.5062 (*), 0.04636	5, 0.07428, 0.08698, 0.11616, 0.12743		
No Best Found for prob = 0.75 No Best Found for prob = 0.75 No Best Found for prob = 0.75 No Best Found for prob = 0.75			
No Best Found for prob = 0.75			
Model Name: Kursawe Searcher Name: MWS Seed: 1 Lives: 3 Al2test: 0.6			
MaxWalkSat Options: Prob: 0.75 MaxTries: 500 MaxChanges 500 Threshold: le-06 Slices: 10			
Time to run (s): 3.537633 No valid runs!			
Model Name: Kursawe Searcher Name: GA Seed: 1 Lives: 3 Al2test: 0.6			
GA Options: popSize: 50 Crossover: 0.6 Gens: [10] Time to run (s): 13.808445			
Runs: 5 Average per run (s): 2.761689			
	1, 0.01386, 0.01645, 0.03094, 0.08662		
Scott-Knott for Kursawe rank , name , med , iqr			
	*	0.02. 0.03.	0.09
1, SA, 9, 5(*	0.09, 0.12,	0.13
Model Name: ZDT1 Searcher Name: SA Seed: 1 Lives: 3 Alltest: 0.6			
SA Options: KMAX: 500 Cooling: 0.6 Time to run (s): 1.384426			
Dung: h			
Runs: 5 Average per run (s): 0.2768852 (*), 0.07102	2, 0.11667, 0.21154, 0.22530, 0.24614		
Average per run (s): 0.2768852 (*), 0.07102			
Average per run (s): 0.2768852 (*), 0.07102 			
Average per run (s): 0.2768852 (*), 0.07102			
Average per run (s): 0.2768852 (*), 0.07102			
Average per run (s): 0.2768852 (*), 0.07102	22, -0.03011, -0.02254, -0.00696, -0.00374	ı	
Average per run (s): 0.2768852 (*	22, -0.03011, -0.02254, -0.00696, -0.00374	i.	
Average per run (s): 0.2768852 (*	22, -0.03011, -0.02254, -0.00696, -0.00374	ı	
Average per run (s): 0.2768852 (*	22, -0.03011, -0.02254, -0.00696, -0.00374	i.	

```
csc710sbse: HW5:Theisen
  Oct 05, 14 21:51
                                                                               Page 3/4
_____
Scott-Knott for ZDT1
rank .
         name , med , igr
         MWS , -2 , 2 (--*
GA , 11 , 4 ( -
SA , 21 , 11 ( ---
                                                ),-0.05, -0.03, -0.02, -0.01, -0.00
- ), 0.09, 0.09, 0.11, 0.13, 0.19
*-- ), 0.07, 0.12, 0.21, 0.23, 0.25
  2 ,
_____
Model Name: ZDT3
Searcher Name: SA
Seed: 1 Lives: 3
Al2test: 0.6
SA Options:
KMAX: 500 Cooling: 0.6
Time to run (s): 1.50074
Runs: 5
Average per run (s): 0.300148
(* | ), -0.03604, 0.21725, 0.23301, 0.25875, 0.29595
Model Name: ZDT3
Searcher Name: MWS
Seed: 1 Lives: 3
A12test: 0.6
MaxWalkSat Options:
Prob: 0.75
MaxTries: 500 MaxChanges 500
Threshold: 1e-06 Slices: 10
Time to run (s): 1.599185
Runs: 5
Average per run (s): 0.319837
(* | ), -0.04109, -0.02913, -0.02180, -0.00308, -0.00266
_____
Model Name: ZDT3
Searcher Name: GA
Seed: 1 Lives: 3
A12test: 0.6
GA Options:
popSize: 50 Crossover: 0.6
Gens: [10]
Time to run (s): 8.908521
Runs: 5
Average per run (s): 1.7817042
(* | ), 0.07100, 0.08388, 0.09071, 0.10566, 0.22608
Scott-Knott for ZDT3
        name , med , iqr
rank .
         1 ,
2 ,
2 ,
 Model Name: Viennet3
Searcher Name: SA
Seed: 1 Lives: 3
Seed I lives. 3
Alletest: 0.6
SA Options:
KMAX: 500 Cooling: 0.6
Time to run (s): 0.221635
Runs: 5
Average per run (s): 0.044327
(* ), 0.00093, 0.00102, 0.00218, 0.00222, 0.00516
No Best Found for prob = 0.75
Model Name: Viennet3
Searcher Name: MWS
Seed: 1 Lives: 3
Al2test: 0.6
MaxWalkSat Options:
Prob: 0.75
MaxTries: 500 MaxChanges 500
Threshold: 1e-06 Slices: 10
Time to run (s): 0.27679
No valid runs!
_____
Model Name: Viennet3
Searcher Name: GA
Seed: 1 Lives: 3
A12test: 0.6
GA Options:
popSize: 50 Crossover: 0.6
Gens: [10]
```

Oct 05, 14 21:51										F	Printed	by C	hristo	pher Th	eise
Runs: 5 Average per run (s): 0.2852062 (*						csc	710sl	bse: I	1W5:	Thei	sen			Page	4/4
rank, name, med, igr 1, GA, 0, 0(-*	Runs: 5 Average per (* ======== Scott-Knott	run (s):	0.28 ===== met3	5206), 0	=====		.=====		.00333,	0.0044	13				
1, GA, 0, 0 (* * 1,0.00,0.00,0.00,0.00,0.00,0.00 1; SA; 0, 0 (* *	rank ,	name ,	med		iqr 										
	1 ,	GA , SA ,	0	,	0 (-	*	*		}	, 0.00,	0.00,	0.00,	0.00,	0.00	

```
csc710sbse: HW5:Theisen
  Oct 05, 14 21:49
                                                                                                                                            Page 1/1
import sys
from datetime import datetime
import random
sys.dont_write_bytecode = True
from models import *
from searchers import *
from utils import *
from options import
from sk import *
myOpt = Options()
#Inspired by vivekaxl's display function
def display(model, searcher, startTime, scores, r):
 print "Model Name: ", model.__name__
print "Searcher Name: ", searcher.__class_.__name__
diff = (datetime.now() - startTime).total_seconds()
  myOpt.printGlobals()
  searcher.printOptions()
print "Time to run (s): ", diff
if r = 0:
  print "No valid runs!"
 def main(modelList, searcherList):
  for klass in modelList:
     classScoreList = []
for searcher in searcherList:
       or searcher in searcherList:
fullScoreList = []
startTime = datetime.now()
scores = []
mySearcher = searcher()
random.seed(myOpt.seed)
        for _ in range(r):
   myKlass = klass()
           result, valid = mySearcher.run(myKlass)
if valid = True:
             scores.append(result)
        display(klass, mySearcher, startTime, scores, len(scores)) fullScoreList.append(searcher.__name__) for x in scores:
        fullScoreList.append(x)
if scores:
     classScoreList.append(fullScoreList)
print "Scott-Knott for", klass.__name__
     print classScoreList
     rdivDemo(classScoreList)
modelList = [Fonseca, Schaffer, Kursawe, ZDT1, ZDT3, Viennet3]
#modelList = [SA, MWS, GA]
#modelList = [Fonseca]
#searcherList = [GA]
main(modelList, searcherList)
```

```
csc710sbse: HW5:Theisen
  Sep 27, 14 18:32
## Hyptotheis Testing Stuff
### Standard Stuff
#### Standard Headers
from future import division
import sys, random, math
sys.dont_write_bytecode = True
#### Standard Utils
class o():
 "Anonymous container"

def __init__(i,**fields) :
    i.override(fields)
  def override(i,d): i.__dict__.update(d); return i
  def __repr__(i):
    d = i.__dict_
   def show(i):
   return [k for k in sorted(i.__dict__.keys())
if ¬ "_" in k]
Misc functions:
rand = random.random
any = random.choice
seed = random.seed
exp = lambda n: math.e**n
     = lambda n: math.log(n,math.e)
     = lambda n: round(n,2)
def median(lst,ordered=False):
 if ¬ ordered: lst= sorted(lst)
 n = len(lst)
 p = n//2
if n % 2: return lst[p]
 q = p - 1
q = max(0,min(q,n))
return (lst[p] + lst[q])/2
def msecs(f):
 import time
  t1 = time.time()
  return (time.time() - t1) * 1000
def pairs(lst):
  "Return all pairs of items i,i+1 from a list."
  last=lst[0]
 for i in lst[1:]:
    yield last,i
     last = i
def xtile(lst,lo=0,hi=100,width=50,
               chops=[0.1 ,0.3,0.5,0.7,0.9],
marks=["-" ," "," ","-"," "],
               bar="|",star="*",show="%3.0f"):
" " "The function _xtile_ takes a list of (possibly) unsorted numbers and presents them as a horizontal
 xtile chart (in ascii format). The default is a
contracted quintile that shows the 10,30,50,70,90 breaks in the data (but this can be
 changed- see the optional flags of the function).
 def pos(p) : return ordered[int(len(lst)*p)]
  def place(x) :
    return int(width*float((x - lo))/(hi - lo+0.00001))
 def pretty(lst) :
    return ','.join([show % x for x in lst])
ordered = sorted(lst)
        = min(lo,ordered[0])
= max(hi,ordered[-1])
 what = [pos(p) for p in chops]
where = [place(n) for n in what]
out = [""] * width
  for one, two in pairs (where):
    for i in range(one, two):
       out[i] = marks[0]
    marks = marks[1:]
  out[int(width/2)]
 out[place(pos(0.5))] = star
return '('+''.join(out) + ")," + pretty(what)
```

```
csc710sbse: HW5:Theisen
  Sep 27, 14 18:32
                                                                                                                   Page 2/6
def tileX():
  import random
  random.seed(1)
  nums = [random.random()**2 for in range(100)]
  print xtile(nums,lo=0,hi=1.0,width=25,show=" %5.2f")
### Standard Accumulator for Numbers
Note the _lt_ method: this accumulator can be sorted by median values
Warning: this accumulator keeps _all_ numbers. Might be better to use
class Num:
  "An Accumulator for numbers"
  def __init__(i,name,inits=[]):
    i.n = i.m2 = i.mu = 0.0
     i.all=[]
     i._median=None
    i.name = name
i.rank = 0
     for x in inits: i.add(x)
  def s(i)
                    : return (i.m2/(i.n - 1))**0.5
  def add(i,x):
    i. median=None
     i.all += [x]
    delta = x - i.mu
     i.mu += delta*1.0/i.n
  i.m2 += delta*(x - i.mu)

def __add__(i,j):
    return Num(i.name + j.name,i.all + j.all)
  def quartiles(i):
   def p(x) : return int(100*g(xs[x]))
    i.median()
    xs = i.all
n = int(len(xs)*0.25)
  return p(n) , p(2*n) , p(3*n) def median(i):
    if ¬ i._median:
    i.all = sorted(i.all)
       i._median=median(i.all)
  return i._median
def __lt__(i,j):
    return i.median() < j.median()</pre>
  def spread(i):
    i.all=sorted(i.all)
     n1=i.n*0.25
     n2=i.n*0.75
    if len(i.all) ≤ 1:
       return 0
     if len(i.all) ≡ 2:
    return i.all[1] - i.all[0] else:
       return i.all[int(n2)] - i.all[int(n1)]
### The A12 Effect Size Test
def al2slow(lst1,lst2):
  "how often is x in lst1 more than y in lst2?"
  more = same = 0.0
  for x in lst1:
    for y in 1st2:
      if x = y : same += 1
elif x > y : more += 1
  x= (more + 0.5*same) / (len(lst1)*len(lst2))
  return x
def a12(lst1,lst2):
   "how often is x in lst1 more than y in lst2?"
  def loop(t,t1,t2):
   while t1.j < t1.n \( \tau \) t2.j < t2.n:
        h1 = t1.l[t1.j]</pre>
       h2 = t2.1[t2.j]
       h3 = t2.1[t2.j+1] if t2.j+1 < t2.n else None
       if h1> h2:
       t1.j += 1; t1.gt += t2.n - t2.j
elif h1 = h2:
         if h3 \wedge h1 > h3 :
            t1.gt += t2.n - t2.j - 1
         t1.j += 1; t1.eq += 1; t2.eq += 1
       else:
t2,t1 = t1,t2
     return t.gt*1.0, t.eq*1.0
  lst1 = sorted(lst1, reverse=True)
  1st2 = sorted(1st2, reverse=True)
  n1 = len(lst1)
n2 = len(lst2)
```

Page 1/6

```
csc710sbse: HW5:Theisen
   Sep 27, 14 18:32
                                                                                                                                                         Page 3/6
         = o(1=1st1, j=0,eq=0,gt=0,n=n1)
= o(1=1st2, j=0,eq=0,gt=0,n=n2)
   gt,eq= loop(t1, t1, t2)
   return gt/(n1*n2) + eq/2/(n1*n2)
  def f1(): return al2slow(11,12)
def f2(): return al2(11,12)
   for n in [100,200,400,800,1600,3200,6400]:
      11 = [rand() for _ in xrange(n)]
12 = [rand() for _ in xrange(n)]
     t1 = msecs(f1)

t2 = msecs(f2)
      print n, g(f1()),g(f2()),int((t1/t2))
 " " " Output:
                a12(slow) tfast / tslow
n a12(fast)
100 0 53
200 0.48
                 0.48
                            26
72
800 0.5
                0.5
                 0.51
1600 0.51
3200 0.49
                 0.49
                              109
6400 0.5
                 0.5
                             244
## Non-Parametric Hypothesis Testing
The following _bootstrap_ method was introduced in
1979 by Bradley Efron at Stanford University. It was inspired by earlier work on the
jackknife.
Improved estimates of the variance were [developed later][efron01].
[efron01]: http://goo.gl/14n8Wf "Bradley Efron A R.J. Tibshirani. An Introduction to the Bootstrap (Chapman & Hall/CRC M
To check if two populations _(y0,z0)_
are different, many times sample with replacement from both to generate _(y1,z1), (y2,z2), (y3,z3)_.. etc.
def sampleWithReplacement(lst):
    "returns a list same size as list"
 def any(n) : return random.uniform(0,n)
def one(lst): return lst[ int(any(len(lst))) ]
return [one(lst) for _ in lst]
Then, for all those samples,
check if some *testStatistic* in the original pair
hold for all the other pairs. If it does more than (say) 99%
of the time, then we are 99% confident in that the
populations are the same.
In such a _bootstrap_ hypothesis test, the *some property
is the difference between the two populations, muted by the joint standard deviation of the populations.
def testStatistic(y,z):
    ""Checks if two means are different, tempered
   by the sample size of 'y' and 'z' " " "
      tmp1 = tmp2 = 0

for y1 in y.all: tmp1 += (y1 - y.mu)**2

for z1 in z.all: tmp2 += (z1 - z.mu)**2
      s1 = (float(tmp1)/(y.n - 1))**0.5
s2 = (float(tmp2)/(z.n - 1))**0.5
      delta = z.mu - y.mu
      if s1+s2:
        delta = delta/((s1/y.n + s2/z.n)**0.5)
      return delta
The rest is just details:
 + Efron advises
 to make the mean of the populations the same (see
 the _yhat,zhat_ stuff shown below).
+ The class _total_ is a just a quick and dirty accumulation class.
+ For more details see [the Efron text][efron01].
def bootstrap(y0,z0,conf=0.01,b=1000):
   " " "The bootstrap hypothesis test from
   p220 to 223 of Efron's book 'An
   introduction to the boostrap. " "
      "quick and dirty data collector"
```

```
csc710sbse: HW5:Theisen
    Sep 27, 14 18:32
                                                                                                                                                                         Page 4/6
       def __init__(i,some=[]):
    i.sum = i.n = i.mu = 0 ; i.all=[]
          for one in some: i.put(one)
       def put(i.x):
           i.all.append(x);
          i.sum +=x; i.n += 1; i.mu = float(i.sum)/i.n
   def __add__(i1,i2): return total(i1.all + i2.all)
y, z = total(y0), total(z0)
   tobs = testStatistic(y,z)
yhat = [y1 - y.mu + x.mu for y1 in y.all]
zhat = [z1 - z.mu + x.mu for z1 in z.all]
   bigger = 0.0
   for i in range(b):
       if testStatistic(total(sampleWithReplacement(yhat)),
                                    total(sampleWithReplacement(zhat))) > tobs:
          higger += 1
  return bigger / b < conf
#### Examples
def _bootstraped():
   def worker(n=1000,
      return n, mul, sigmal, mu2, sigma2,
   'different' if bootstrap(x,y) else 'same' # very different means, same std
   print worker(mul=10, sigmal=10,
   mu2=100, sigma2=10)
# similar means and std
   print worker(mul= 10.1, sigmal=1,
mu2= 10.2, sigma2=1)
# slightly different means, same std
   print worker(mul= 10.1, sigmal= 1,
   # different in mu eater by large std

print worker(mul= 10.1, sigma1= 10,
                         mu2= 10.8, sigma2= 1)
Output:
 bootstraped()
 (1000, 10, 10, 100, 10, 'different')
 (1000, 10.1, 1, 10.2, 1, 'same')
(1000, 10.1, 1, 10.8, 1, 'different')
 (1000, 10.1, 10, 10.8, 1, 'same')
Warning— the above took 8 seconds to generate since we used 1000 bootstraps. As to how many bootstraps are enough, that depends on the data. There are
results saying 200 to 400 are enough but, since I am suspicious man, I run it for 1000.
Which means the runtimes associated with bootstrapping is a significant issue.
To reduce that runtime, I avoid things like an all-pairs comparison of all treatments (see below: Scott-knott). Also, BEFORE I do the boostrap, I first run
the effect size test (and only go to bootstrapping in effect size passes:
def different(11,12):
   #return bootstrap(11,12) and a12(12,11)
   return a12(12,11) A bootstrap(11,12)
## Saner Hypothesis Testing
The following code, which you should use verbatim does the following:
 + All treatments are clustered into _ranks_. In practice, dozens of treatments end up generating just a handful of ranks. + The numbers of calls to the hypothesis tests are minimized:
   + Treatments are sorted by their median value.
  + Treatments are divided into two groups such that the expected value of the mean values _after_ the split is minimized; 
+ Hypothesis tests are called to test if the two groups are truly difference. 
+ All hypothesis tests are non-parametric and include (1) effect size tests and (2) tests for statistically significant numbers;
        + Slow bootstraps are executed if the faster _A12_ tests are passed:
In practice, this means that the hypothesis tests (with confidence of say, 95%)
 + With this method, 16 treatments can be studied using less than _∑<sub>1,2,4,8,16</sub>log<sub>2</sub>i = 15_hypothesis tests and confidence _0.99<sup + But if did this with the 120 all—pairs comparisons of the 16 treatments, we would have total confidence _0.99<sup>120</sup>=0.30.
```

Page 6/6

Sep 27, 14 18:32 csc710sbse: HW5:Theisen

```
For examples on using this code, see _rdivDemo_ (below)
def scottknott(data,cohen=0.3,small=3, useA12=False,epsilon=0.01):
    """Recursively split data, maximizing delta of
 the expected value of the mean before and
 after the splits.
 Reject splits with under 3 items " " "
  all = reduce(lambda x,y:x+y,data)
   same = lambda l,r: abs(l.median() - r.median()) \le all.s()*cohen
  same = lambda l, r: ¬ different(l.all,r.all)
big = lambda n: n > small
  return rdiv(data,all,minMu,biq,same,epsilon)
def rdiv(data, # a list of class Nums
      all, # all the data combined into one num
      div, # function: find the best split
            big, # function: rejects small splits
            same, # function: rejects similar splits
epsilon): # small enough to split two parts
 " " "Looks for ways to split sorted data,
Recurses into each split. Assigns a 'rank' number
 to all the leaf splits found in this way.
  def recurse(parts.all.rank=0):
      "Split, then recurse on each part."
      cut,left,right = maybeIgnore(div(parts,all,big,epsilon),
                                             same, parts)
        rank = recurse(parts[:cut],left,rank) + 1
rank = recurse(parts[cut],left,rank)
        # if no cut, then all get same rank
        for part in parts:
          part.rank = rank
     return rank
  recurse(sorted(data),all)
  return data
def maybeIgnore((cut,left,right), same,parts):
     if same(sum(parts[:cut],Num('upto'))
        sum(parts[cut:],Num('above'))):
cut = left = right = None
  return cut, left, right
def minMu(parts,all,big,epsilon):
 " " Find a cut in the parts that maximizes
 the expected value of the difference in
 the mean before and after the cut.
 Reject splits that are insignificantly
 different or that generate very small subsets.
  cut,left,right = None,None,None
  before, mu = 0, all.mu
for i,l,r in leftRight(parts,epsilon):
     if big(1.n) ^ big(r.n):
        n = all.n * 1.0

now = l.n/n*(mu-l.mu)**2 + r.n/n*(mu-r.mu)**2
        if now > before:
          before,cut,left,right = now,i,l,r
  return cut, left, right
def leftRight(parts,epsilon=0.01):
 " " Iterator. For all items in 'parts',
return everything to the left and everything
 from here to the end. For reasons of
 efficiency, take a first pass over the data
to pre-compute and cache right-hand-sides
 rights = {}
n = j = len(parts) - 1
while j > 0:
   rights[j] = parts[j]
    if j < n: rights[j] += rights[j+1]
j -=1</pre>
  left = parts[0]
  for i, one in enumerate(parts):
   if i> 0:
        if parts[i]._median - parts[i-1]._median > epsilon:
       yield i,left,rights[i]
left += one
## Putting it All Together
Driver for the demos-
def rdivDemo(data):
  def z(x):
    return int(100 * (x - lo) / (hi - lo + 0.00001))
  data = map(lambda lst:Num(lst[0],lst[1:]),
```

Sep 27, 14 18:32 **csc710sbse: HW5:Theisen**

Page 5/6

```
csc710sbse: HW5:Theisen
    Sep 23, 14 1:52
                                                                                                                                                                                                                 Page 1/1
#From Class Discussion 8/26/2014
from _future__ import division
import sys,re,random,math
import numpy as np
sys.dont_write_bytecode = True
 from options import *
 #Taken verbatim from the class website.
 def pairs(lst):
     last=lst[0]
    for i in lst[1:]:
        yield last,i
last = i
contracted _quintile_ that shows the
  10,30,50,70,90 breaks in the data (but this can be changed—see the optional flags of the function).
    def pos(p): return ordered[int(len(lst)*p)]
def place(x):
    return int(width*float((x - lo))/(hi - lo))
    def pretty(lst):
    return ','.join([show % x for x in lst])
ordered = sorted(lst)
  ordered = sorted(lst)
lo = min(lo,ordered[0])
hi = max(hi,ordered[-1])
what = [pos(p) for p in chops]
where = [place(n) for n in what]
out = [""] * width
for one, two in pairs(where):
    for i in range(one, two):
        out[i] = marks[0]
    marks = marks[1:]
out[int(width/2)] = bar
out[place(pos(0.5))] = star
return ''.join(out) + "." + pretty(what)
```

Sep 29, 14 20:29	csc710sbse: HW5:Theisen	Page 1/1
<pre>from sim_anneal import * from max_walk_sat import * from genetic_alg import *</pre>		
Trom generic_dig Import		
<u> </u>		

```
csc710sbse: HW5:Theisen
   Oct 05, 14 21:36
                                                                                                                                               Page 1/1
#Structure from SA Lecture
import sys,re,random,math
import copy
sys.dont_write_bytecode = True
from options import *
from utils import *
from analyzer import *
myOpt = Options()
#Structure from:
#http://www.cleveralgorithms.com/nature-inspired/evolution/genetic_algorithm.html
class GA:
  def Mutate(self, child, pMut):
   for x in range(0, len(child.XVar)):
     if(random.random() < pMut):</pre>
           child.specificMutate(x)
      return child
  def Crossover(self, parent1, parent2, crossover):
    #BECAUSE REASONS OK TODO
      return parent1, parent2
  def SelectParents(self, pop): #all possible parents
     temp = []
      for x in pop:
        for y in pop:
  if x ≠ y:
              temp.append([x, y])
      return temp
   def GetBestSolutions(self, pop):
     eMin = 1
temp = []
     temp = []
temp2 = []
seen = []
for x in pop:
    if x.Energy() ¬ in seen:
        temp.append([x.Energy(), x])
        seen.append(x.Energy())
temp.sort(key = lambda x: x[0])
for x in range (0, 50):
    temp2.aopend(temp[x][1])
        temp2.append(temp[x][1])
#print temp[x]
      return temp[0][0], temp2
   def run(self, klass):
     ga = klass
pMutate = 1.0/len(ga.XVar)
XVarBest = ga.XVar
#print 'start energy: ', eBest
      k = 1
eList = []
      population = []
      analyze = Analyzer()
stop = False
      for i in range(myOpt.ga_pop_size):
        population.append(copy.deepcopy(ga)) #add a randomly generated model to list
      for gens in myOpt.ga_gen_list:
    eBest = 1
    while k < gens:</pre>
           parents = self.SelectParents(population) #all possible pairings of parents children = [] for parent1, parent2 in parents:
              child1, child2 = self.Crossover(parent1, parent2, myOpt.ga_crossover)
children.append(self.Mutate(child1, pMutate))
children.append(self.Mutate(child2, pMutate))
            eBest, population = self.GetBestSolutions(children)
            eList.append(eBest)
        #some "is significantly better" termination logic here
      return min(eList), True
   def printOptions(self):
     print "popSize:", myOpt.ga_pop_size, "Crossover:", myOpt.ga_crossover
print "Gens:", myOpt.ga_gen_list
```

csc710sbse: HW5:Theisen Oct 05, 14 21:40 Page 1/1 #Structure from SA Lecture import sys,re,random,math sys.dont_write_bytecode = True from options import * from utils import * from analyzer import myOpt = Options() class MWS: debug = False def say(self, x): sys.stdout.write(str(x)); sys.stdout.flush() def specificRun(self, probability, klass): fon = klass XVarBest = fon.XVar eBest = e = 1 eNew = 1 k = 1 temp = [] self.say(int(math.fabs(eBest-1)*100)) self.say('') analyze = Analyzer() stop = False for i in xrange(myOpt.mws_maxTries): fon.Chaos() for j in xrange(myOpt.mws_maxChanges): eNew = fon.Energy() if(eNew < myOpt.mws_threshold v stop = True): #% means found a solution and quit</pre> self.say('%') eBest = eNew XVarBest = list(fon.XVar) temp.append(eNew) #print xtile(temp,lo=0, hi=1, width=25,show=" %1.5f") return eBest, XVarBest #modify random part of solution if probability < random.uniform(0,1): fon.Neighbor()</pre> self.say('+') #maximize for some random else: fon.BestNeighbor() self.say('.') temp.append(eNew) if (i+1)*(j+1) % 40 ≡ 0 ∧ len(temp) ≠ 0: #print '' self.say(int(math.fabs(eNew-1)*100)) #print xtile(temp,lo=0, hi=1, width=25,show=" %1.5f") #stop = analyze.EraStop(temp) temp = [] return -1, XVarBest def run(self, klass): theBest = -1 valid = False eBest, XVarBest = self.specificRun(myOpt.mws_prob, klass) if eBest = -1: print 'No Best Found for prob = ', myOpt.mws_prob self.say('' else: theBest = eBest valid = True return theBest, valid def printOptions(self): print "MaxWalkSat Options:" print "MaxWalkSat Options:" print "Prob:", myOpt.mws_prob print "MaxTries:", myOpt.mws_maxTries, "MaxChanges", myOpt.mws_maxChanges print "Threshold:", myOpt.mws_threshold, "Slices:", myOpt.mws_slices

csc710sbse: HW5:Theisen Oct 05, 14 21:41 Page 1/1 #Structure from SA Lecture import sys,re,random,math sys.dont_write_bytecode = True from options import * from utils import * from analyzer import myOpt = Options() class SA: def say(self, x): if myOpt.debug: sys.stdout.write(str(x)); sys.stdout.flush() def run(self, klass): XVarBest = sa.XVar eBest = e = 1 #print 'start energy: ', eBest temp = [] self.say(int(math.fabs(eBest-1)*100)) self.say('') analyze = Analyzer() stop = False while k < myOpt.sa_kmax \(\) stop \(\) False: sa.Neighbor() eNew = sa.Energy()</pre> if eNew < eBest: eBest = eNew XVarBest = list(sa.XVar) self.say('!') if eNew < e:</pre> e = eNew self.say('+') self.say('+') #Probability Check from SA Lecture elif math.exp(-1*(eNew-e)/(k/myOpt.sa_kmax**myOpt.sa_cooling))<random.uniform(0,1): #P function should be between 0 and 1 #more random hops early, then decreasing as time goes on sa.Chaos() self.say('?') self.say('.') k = K + T temp.append(eBest) if k % 50 = 0 ^ k ≠ myOpt.sa_kmax ^ len(temp) ≠ 0: self.say(int(math.fabs(eBest-1)*100)) self.say('' #print xtile(temp,lo=0, hi=1,width=25,show=" %1.5f") stop = analyze.EraStop(temp) temp = [] if myOpt.debug: #print '\nFound best - e: ', eBest #print 'Variables: ' for vars in XVarBest: self.say(vars) self.say(",") #print "\n" return eBest, True def printOptions(self): print "SA Options:" print "KMAX:", myOpt.sa_kmax, "Cooling:", myOpt.sa_cooling

Sep 29, 14 21:48	csc710sbse: HW5:Theisen	Page 1/1
<pre>from fonseca_model import * from schaffer_model import * from kursawe_model import * from ZDT1_model import * from ZDT3_model import * from viennet3_model import *</pre>		
<pre>from DTLZ7_model import *</pre>		


```
csc710sbse: HW5:Theisen
  Oct 05, 14 21:36
                                                                                                                          Page 1/1
#From Class Discussion 8/26/2014
from _future__ import division
import sys,re,random,math
import numpy as np
sys.dont_write_bytecode = True
from options import *
myOpt = Options()
rand = random.random
class Model:
  #Default Values overwritten by subclass; should have better defaults, but...
  smin = 1
  smax = 1

XVar = [random.uniform(smin, smax) for i in range (0, n)]

XVarMax = XVar
  eMax = 0
eMin = 0
  def Energy(self):
    raise NotImplementedError()
  def RawEnergy(self):
     raise NotImplementedError()
  def Neighbor(self):
     self.XVar[random.randint(0, self.n-1)] = random.uniform(self.smin, self.smax)
  def specificMutate(self, x):
     self.XVar[x] = random.uniform(self.smin, self.smax)
  def BestNeighbor(self):
     toChange = random.randint(0, self.n-1)
     toIncrement = (self.smax - self.smin) / myOpt.mws_slices
curMax = 1
     maxVal = self.XVar[toChange]
     for i in xrange(myOpt.mws_slices):
    self.XVar[toChange] = self.smin + toIncrement
       x = self.Energy()
       if x < curMax:</pre>
         curMax = x
maxVal = self.XVar[toChange]
  def Reset(self):
    self.XVar = self.XVarMax
  def Chaos(self):
     for vars in self.XVar:
       vars = random.uniform(self.smin, self.smax)
  def Baseline(self, numRuns):
     self.Chaos()
     self.eMax = self.eMin = self.RawEnergy()
     runs = 1
     while runs < numRuns:
       self.Neighbor()
eNew = self.RawEnergy()
if eNew > self.eMax: #find largest difference
         self.eMax = eNew
self.XVarMax = self.XVar
          #print self.XVarMax, eNew
       if eNew < self.eMin: #find smallest difference
    self.eMin = eNew</pre>
          #print 'Min: ', self.XVar, eNew
     runs += 1
#print 'Baseline: ', self.eMin, ', ', self.eMax
  def __init__(self):
     raise NotImplementedError()
```


csc710sbse: HW5:Theisen Page 1/1 Oct 05, 14 20:29 #From Class Discussion 8/26/2014 from _future__ import division import sys,re,random,math import numpy as np sys.dont_write_bytecode = True from model_base import * from options import * class Viennet3(Model): smin = -3.0 smax = 3 n = 2 XVar = [random.uniform(smin, smax) for i in range (0, n)] XVarMax = XVar eMax = 0 eMin = 0 def Energy(self): Ner Energy(seif): X = self_XVar f1 = 0.5*X[0]**2 + X[1]**2 + math.sin(X[0]**2+X[1]**2) f2 = (3*X[0]-2*X[1]+4)**2/8 + (X[0]-X[1]+1)**2/27 + 15 f3 = 1/(X[0]+X[1]+1) - 1.1*math.e**(-X[0]**2-X[1]**2) return (math.fabs(f1+f2+f3) - self.eMin) / (self.eMax - self.eMin) def RawEnergy(self): lef RawEnergy(self): X = self.XVar fl = 0.5*X[0]**2 + X[1]**2 + math.sin(X[0]**2+X[1]**2) f2 = (3*X[0]-2*X[1]+4)**2/8 + (X[0]-X[1]+1)**2/27 + 15 f3 = 1/(X[0]+X[1]+1) - 1.1*math.e**(-X[0]**2-X[1]**2) return math.fabs(f1+f2+f3) def __init__(self): self.Baseline(10000) self.XVar = self.XVarMax