decrypt

Ashwath Raj

May 10, 2021

Report:

I set up an energy function based on the pairwise frequencies. I use a for loop to show all the energies given substitution shifts from 0-25. I saw all the output and I don't see an immediate "solve" or a point where manual tweaks made sense.

I haven't implemented mcmc() yet, but I'm curious how it would change the result because it looks like possible states are the possible shifts I already tested.

```
# read coded message
input <- readLines("CodedMessage_Short.txt", n=1000)
input2 <- readLines("CodedMessage_Med.txt", n=1000)
# read frequency table
p <-read.table("LetterPairFreqFrom7Novels.txt")
input <- tolower(input)
s <- strsplit(input, split = " ")
s <- do.call(paste0, as.data.frame(s))
s2 <- strsplit(input2, split = " ")
s2 <- do.call(paste0, as.data.frame(s2))

# p[,1]
#
# for (word in s){
#     for(char in 1:str_length(word)){
#       print(char)
#     }
# }

#At a given substitution point, evaluate the energy
energy <- function(strings, pairWeight, shift){
  freqs = 0
  alphabet <- letters#[1:26]
  wls <- str_length(strings) #word lengths

  for (word in strings){
    for(char in 1:(str_length(word)-1)){
        # if (char < str_length(word)){ #Can't index past the last letter
         letter1 = (match(substr(word,char,char),alphabet) + shift) %% 26
         letter2 = (match(substr(word,char+1,char+1), alphabet) + shift) %% 26

         if(!is.na(letter1) & !is.na(letter2)){
           if(letter1 != 0 & letter2 != 0){
              freqs = freqs + (pairWeight[letter1,letter2])
              # print(freqs)
         }
      }
                # for (line in p){ #Every line of frequencies
        #   for (letter in 1:26){ #Every letter frequency
        #
        #   }
        # }
    }
    # browser()
  }
  return(freqs)

}

#At a given substitution point, evaluate the energy
convert <- function(strings, shift){
  new_msg = ""
  for (word in strings){
    new_word = ""
      for(char in 1:(str_length(word))){
        # chars = paste(char,())
        letter_num = (match(substr(word, char, char),letters) + shift)
        if(!is.na(letter_num)){
          if(letter_num != 0){
          if (letter_num > 26){letter_num = letter_num %% 26}
          letter = letters[letter_num]
          new_word = paste(new_word, letter, sep = "")
          }
        }
      }
    new_msg = paste(new_msg, new_word, sep = " ")
  }
  return(new_msg)
}


#Finding the best energy for a shift for s
max = 0
best_shift = 0
for (shift_x in 0:(length(letters)-1)){
  fx = energy(s,p,shift_x)
  print(shift_x)
  print(fx)
  # print(convert(s,shift_x))
  if(fx > max){
    max = fx
    best_shift = shift_x
   }
}
```

```
## [1] 0
## [1] 19968.4
## [1] 1
## [1] 18488.5
## [1] 2
## [1] 12131.9
## [1] 3
## [1] 10773.8
## [1] 4
## [1] 8379.3
## [1] 5
## [1] 14704.9
## [1] 6
## [1] 10883.6
## [1] 7
## [1] 16792.1
## [1] 8
## [1] 9810.6
## [1] 9
## [1] 8476.7
## [1] 10
## [1] 11702.2
## [1] 11
## [1] 14991.1
## [1] 12
## [1] 18224.8
## [1] 13
## [1] 12402.1
## [1] 14
## [1] 13318.9
## [1] 15
## [1] 19134
## [1] 16
## [1] 9955.5
## [1] 17
## [1] 5903.9
## [1] 18
## [1] 7768.2
## [1] 19
## [1] 20146.9
## [1] 20
## [1] 18888.2
## [1] 21
## [1] 6935.7
## [1] 22
## [1] 10661.9
## [1] 23
## [1] 12873.6
## [1] 24
## [1] 18084.9
## [1] 25
## [1] 11794.9
```

```
print(convert(s,shift = best_shift))
```

```
## [1] " zsh nveyz zvih v javl hdhy mx zheed jhxxmu sh kay leoxf vx a emjjyemdbh yvjqhe keavzs mozyvlh zsh zheeabh mn zsh la
xbhey zsh taefvxw jmz azzhxlaxz sal remowsz zsh bae moz axl sh kay yzvjj smjlvxw zsh lmme mthx rhbaoyh zheed jhxxmuy jhnz nm
mz kay yzvjj laxwjvxw mozyvlh ay vn sh sal nmewmzzhx sh sal mxh sh sal a dmoxwjmmfvxw nabh roz svy save kay rmxh ksvzh dmo b
mojl zhjj rd svy hdhy zsaz sh kay tjayzhehl zm zsh savejvxh roz mzshekvyh sh jmmfhl jvfh axd mzshe xvbh dmoxw wod vx a lvxxh
e cabfhz ksm sal rhhx ythxlvxw zmm iobs imxhd vx a cmvxz zsaz huvyzy nme zsaz toetmyh axl nme xm mzshe zsheh kay a wvej rhyv
lh svi she save kay a jmqhjd ysalh mn laef ehl axl ysh sal a lvyzaxz yivjh mx she jvty axl mqhe she ysmojlhey ysh sal a rjoh
ivxf zsaz ajimyz ialh zsh emjjyemdbh jmmf jvfh coyz axmzshe aozmimrvjh vz lvlxz povzh xmzsvxw bax zsh azzhxlaxz kay zsh oyoa
j sajnzmows bsaeabzhe vx a ksvzh bmaz kvzs zsh xaih mn zsh ehyzaoeaxz yzvzbshl abemyy zsh nemxz mn vz vx ehl sh kay whzzvxw
nhl ot jmmf ivyzhe sh yavl kvzs ax hlwh zm svy qmvbh kmojl dmo ivxl a ksmjh jmz tojjvxw dmoe jhw vxzm zsh bae ym v bax fvxl
mn ysoz zsh lmme me ysmojl v mthx vz ajj zsh kad ym dmo bax najj moz zsh wvej waqh svi a jmmf ksvbs mowsz zm saqh yzobf az j
hayz nmoe vxbshy moz mn svy rabf vz lvlxz rmzshe svi hxmows zm wvqh svi zsh ysafhy az zsh laxbhey zshd whz zsh ymez mn thmtj
h zsaz lvyvjjoyvmx dmo armoz ksaz a jmz mn wmjnvxw imxhd bax lm nme zsh theymxajvzd"
```

```
#Finding the best energy for a shift for s2
max = 0
best_shift = 0
for (shift_x in 0:(length(letters)-1)){
  fx = energy(s,p,shift_x)
  print(shift_x)
  print(fx)
  # print(convert(s,shift_x))
  if(fx > max){
    max = fx
    best_shift = shift_x
   }
}
```

```
## [1] 0
## [1] 19968.4
## [1] 1
## [1] 18488.5
## [1] 2
## [1] 12131.9
## [1] 3
## [1] 10773.8
## [1] 4
## [1] 8379.3
## [1] 5
## [1] 14704.9
## [1] 6
## [1] 10883.6
## [1] 7
## [1] 16792.1
## [1] 8
## [1] 9810.6
## [1] 9
## [1] 8476.7
## [1] 10
## [1] 11702.2
## [1] 11
## [1] 14991.1
## [1] 12
## [1] 18224.8
## [1] 13
## [1] 12402.1
## [1] 14
## [1] 13318.9
## [1] 15
## [1] 19134
## [1] 16
## [1] 9955.5
## [1] 17
## [1] 5903.9
## [1] 18
## [1] 7768.2
## [1] 19
## [1] 20146.9
## [1] 20
## [1] 18888.2
## [1] 21
## [1] 6935.7
## [1] 22
## [1] 10661.9
## [1] 23
## [1] 12873.6
## [1] 24
## [1] 18084.9
## [1] 25
## [1] 11794.9
```

```
print(convert(s2,shift = best_shift))
```

```
## [1] " sh nveyz zvih  javl hdhy mx heed hxxmu sh kay leoxf vx a mjjymdbh vjqhe eavzs mozyvlh zsh zheeabh mn sh axbhey sh t
aefvxw jmz azzhxlaxz sal remowsz zsh bae moz axl sh kay yzvjj smjlvxw zsh lmme mthx rhbaoyh heed hxxmuy jhnz nmmz kay yzvjj
laxwjvxw mozyvlh ay vn sh sal nmewmzzhx sh sal mxh h sal a dmoxwjmmfvxw nabh roz svy save kay rmxh ksvzh mo bmojl zhjj rd sv
y hdhy zsaz sh kay tjayzhehl zm zsh savejvxh roz mzshekvyh sh jmmfhl jvfh axd mzshe xvbh dmoxw wod vx a lvxxhe cabfhz ksm sa
l rhhx ythxlvxw zmm iobs imxhd vx a cmvxz zsaz huvyzy nme zsaz toetmyh axl nme xm mzshe sheh kay a wvej rhyvlh svi he save k
ay a jmqhjd ysalh mn laef ehl axl ysh sal a lvyzaxz yivjh mx she jvty axl mqhe she ysmojlhey ysh sal a rjoh ivxf zsaz ajimyz
ialh zsh mjjymdbh jmmf jvfh coyz axmzshe aozmimrvjh z lvlxz povzh mzsvxw bax sh azzhxlaxz kay zsh oyoaj sajnzmows bsaeabzhe
vx a ksvzh bmaz kvzs zsh xaih mn zsh ehyzaoeaxz yzvzbshl abemyy zsh nemxz mn vz vx ehl h kay whzzvxw nhl ot mmf ivyzhe sh ya
vl kvzs ax hlwh zm svy qmvbh kmojl dmo ivxl a ksmjh jmz tojjvxw dmoe jhw vxzm zsh bae ym  bax fvxl mn ysoz zsh lmme e ysmojl
mthx vz ajj zsh kad ym dmo bax najj moz sh wvej waqh svi a jmmf ksvbs mowsz zm saqh yzobf az jhayz nmoe vxbshy moz mn svy ra
bf z lvlxz rmzshe svi hxmows zm wvqh svi zsh ysafhy z sh axbhey zshd whz zsh ymez mn thmtjh zsaz lvyvjjoyvmx dmo armoz ksaz
a jmz mn wmjnvxw imxhd bax lm nme zsh theymxajvzd  jmkykoxw nmehvwx ythhlyzhe kvzs xm zmt levnzhl vxzm zsh taefvxw jmz axl a
iax wmz moz mn vz axl oyhl zsh lays jvwszhe mx a jmxw bvwaehzzh h kay khaevxw a tojjmqhe bshbf ysvez dhjjmk yjabfy axl evlvx
w rmmzy h yzemjjhl mnn zeavjvxw bjmoly mn vxbhxyh xmz hqhx rmzshevxw zm jmmf zmkaely zsh mjjymdbh h temrarjd zsmowsz vz kay
bmexd z zsh nmmz mn zsh yzhty ot zm zsh zheeabh sh taoyhl zm yzvbf a imxmbjh vx svy hdh sh wvej yavl kvzs a xvbh roeyz mn bs
aei  saqh a kmxlhenoj vlha laejvxw sd lmxz kh coyz zafh a bar zm dmoe tjabh axl whz dmoe bmxqhezvrjh moz zy yobs a kmxlhenoj
xvwsz nme a eox ot zsh bmayz zm mxzhbvzm  fxmk ymih thmtjh zsheh ksm aeh zsemkvxw a laxbh aemoxl zsh tmmj sh ksvzhsavehl jal
yavl tmjvzhjd knojjd ymeed roz  lmxz saqh vz axd imeh  kay bmithjjhl zm yhjj vz emi svy qmvbh axl aezvbojazvmx dmo kmojlxz s
aqh fxmkx sh sal sal axdzsvxw yzemxwhe zsax meaxwh covbh zm levxf mjl vz laejvxw mk lm dmo ihax sh yjvl akad nemi svi ajmxw
zsh yhaz roz she qmvbh yjvl akad a jmz naezshe zsax zsaz  ihax  sal zm sh yavl me hazvxw imxhd s  yhh  yjvbh mn ytoimxv kmoj
lxz saqh ihjzhl mx she xmk sh azzhxlaxz sal zsh ksvzhsavehl rmd evwsz ksheh sh bmojl ehabs svi  vx a jmkvxbmih reabfhz mmf r
oyzhe sh yavl qh wmz zm toz a bae akad hh dmo ymih imeh ymih mzshe zvihiadrh h jhz zsh lmme ykvxw mthx sh leoxf temitzjd yjv
l mnn zsh yhaz axl jaxlhl mx zsh rjabfzmt mx zsh yhaz mn svy taxzy m  khxz mqhe axl lemtthl id xvbfhj  wohyy vzy ajkady a iv
yzafh zm vxzhenheh kvzs a leoxf qhx vn sh fxmky axl jvfhy dmo sh vy ajkady jvarjh zm saoj mnn axl tmfh dmo vx zsh zhhzs  wmz
svi oxlhe zsh aeiy axl wmz svi ot mx svy nhhz saxf dmo ym qhed iobs sh yavl tmjvzhjd sh wvej yjvl oxlhe zsh kshhj h whzy ym
wmllai xwjvys kshx shy jmalhl ysh yavl vx a yzavxjhyyyzhhj qmvbh saxfy nme bazbsvxw svi jj whz svi vx zsh rabf mn zsh bae  y
avl i zheevrjd ymeed i jazh nme ax hxwawhihxz sh jhz zsh bjozbs vx axl zsh mjjy yzaezhl zm wjvlh hy coyz a jmyz lmw ysh allh
l kvzs a bmmj yivjh hesaty dmo bax nvxl a smih nme svi hy smoyhremfhx  imeh me jhyy xl zsh mjjy zvbfhl lmkx zsh hxzeaxbh lev
qhkad mxzm oxyhz mojhqael ialh a evwsz zoex axl kay wmxh  kay jmmfvxw anzhe she kshx zsh azzhxlaxz baih rabf xl  kay yzvjj s
mjlvxw zsh iax ot axl sh kay xmk ymoxl ayjhht hjj zsazy mxh kad mn lmvxw vz  zmjl zsh ksvzh bmaz oeh  sh yavl bdxvbajjd sd k
ayzh vz mx a joys shi boeqhy axl ajj mo fxmk svi  shael zsh laih bajj svi heed zshekvyh  lmxz fxmk svi nemi a bmky barmmyh o
z  mxjd rhhx sheh zkm khhfy hz id bae kvjj dmo  waqh svi zsh zvbfhz d zsh zvih sh remowsz id jly mqhe  nhjz ay vn  kay smjlv
xw ot a yabf mn jhal sh ksvzh bmaz shjthl ih whz svi vxzm zsh nemxz yhaz sh boyzmihe mthxhl ax hdh axl zsaxfhl oy axl khxz z
m yjhht awavx hy zsh tmjvzhyz leoxf  hqhe ihz  yavl zm zsh ksvzh bmaz shd bmih ajj yvghy axl ysathy axl ajj fvxly mn iaxxhey
sh yavl xl zshdeh ajj roiy mmfy jvfh zsvy mxh sal a tjayzvb cmr mxh zvih has  waqh svi a lmjjae axl sh zsaxfhl ih h kay evws
z armoz zsh tjayzvb cmr sh evwsz yvlh mn id xhk nevhxly nabh kay nemghx axl ksvzvys axl yhaihl kvzs zsvx nvxh ybaey sh yfvx
sal a wjmyyd jmmf ajmxw zsh ybaey  tjayzvb cmr axl a tehzzd leayzvb mxh sazbsa avi zm lm kvzs svi afh svi smih axl ymrhe svi
ot hxmows zm zhjj ih ksheh sh jvqhy sh ksvzh bmaz wevxxhl az ih fad yobfhe n vz kay ih l coyz lemt svi vx zsh wozzhe axl fhh
t wmvxw shi rmmgh smoxly coyz iafh a iax a jmz mn zemorjh nme xm nox  wmz a tsvjmymtsd armoz zshi zsvxwy sh kad zsh bmithzvz
vmx vy xmkalady a wod say zm yaqh svy yzehxwzs zm temzhbz svyyhjn vx zsh bjvxbshy  bax yhh dmoqh ialh a rvw yobbhyy moz mn v
```

```
z  yavl h jmmfhl toggjhl axl zshx sh yzaezhl zm whz ial roz rd zsaz zvih  kay vx zsh bae axl imqvxw h kay taezjd evwsz mn bm
oeyh heed hxxmu ialh ih tjhxzd mn zemorjh oz anzhe ajj zsazy id jvxh mn kmef"
```