Assignment 1

Andrew Rosen

September 7, 2014

Question 0

There are a couple of ways this could work. If we see substitution of Chinese characters for other Chinese characters, we are looking at a frequency table with an enormous x axis. This just means we need a larger corpus to use brute force method.

On the other hand, Chinese is written down phonetically, it would be a frequency analysis on syllables, which is also a larger space.

1

2

No, there aren't enough letters.

3

We can discount other cipher systems because it's unlikely that an othery cipher system would yield a ciphertext consisting only of `adfgvx`.

It's almost certainly permuted, as $\Phi = 0.00539088589576$ on the corpus broken up into pairs of symbols, which looks just like random text. If it wasn't permuted,Φ would be much closer to 0.0385.

4

A Φ of 0.0363075549455 was found with 115 columns. This was significantly larger than the other Φ values. The resulting text:

```
   vxvdaagvaffvxvddddxavafvggvvdaagvvxaadfvggggvggvgfdxvdxgaaavg
gvdxxfxfddavggdgdgdfdxaaavxfafdxfvgvafvgddaavggadfafvdaaaxdxfadxd
xvgaadgaaavgvdfavgvggafgvaafvaafaggdfgvvddxvaaffvdxaafvdxvdaaavva
xfdfdxaxdxavxfdxddgvdgdxddggavaavaaadfgvvddxdfvaafvavggggvfxvgggv
xgvvddxvggvvdaagvafgvvxaafvdgdfxfafaxdxavdgdxvaafgaaaxffvgaafdxvg
```

```
gadxvxaafvafvgaaddaaavxfdxfvfvaavaaxaavggadxvafvgvaagvdxgvvdaavgv
gggvxaavgvagvvddxdfxvfvdxvagvggfgxvgvafgvvaggvxvggvggxfaaxaafvgdx
fvfvvxvddfdfggxvfvfxxvxffxafvgfdxfafgvgvxfdxvadxaxafxfdfggxvgvvdd
xdfvxggxvxfvafvaadffddxgvxvfgaavgvavaggfvggdgdxgvvdafvgfdddggavdf
ggxvavxfafaxafvgfdgaaavggvdfggxvvggggvfxvgggvxafvgfdggddgaggxvavf
vdxgvvdaagvafvxaafvafxfxfaavgvagvvddxdfvaafvavggvfdafaxdxdgdxfgaf
xfxffvgvvddxdffdaaaxdxdgdxgaxfxvdgfgfvggvggvvddxfvafvadxggddgvvdd
xvddxaavaaavgvafvgvavaavgfddxaafvafgvdgaadfaafgfgdxaaavgvvdggfvdx
gaxfxvdgfgfvggvggvvddxvddxaavaggddgvdxvggaxvavdxvadgdxddggavgvvdd
xgvafdgdxfadxafvgfdafvdaaaxdxfxvgggvxvgggvgdxgaxfxvdgfgggvggvvddx
vddxaavavdaaaxdxdgggavdxdxddddxgagvxvfgggvgdgdfxfafaxdxavaavgvad
gaafxdxdgdxdddxdxxfdgggavdxaavgffafggxvfvgvggfdggfvgvavaaaffdvdgv
aavxaadfgvvddxvgaavgvagvvddxavdxaavgvavaggvxvdaagvvxaafvvxaavggvd
xvagvggfadxvaggvgdxvxafgvvdggxvgvddxvavgvvddxavxfggfvfvggddgvafdg
dxgvvdaavgaavxvdggxfdxfaggffggddfgafxfxffvvaggdxfvvgggvxdfggxvfxv
gggvxafgvggddgvdxvgaffvfvgggvvdggfvdxfvafdgfgxfdxggxfvaddaafvvdaf
ggvgdxvaavdxdgdxvaafdxfvaaavdxfvggdgdxgvafdgdxfvdgggavdxdxddddafg
aaagaafggxvfvgvvdaavgaaxfxfgvvddxvaaffvfgdxvgfvaaavdffvgvxvdddd
```

5

I decided to make a crib of the first 10 occurrences of the symbol gg, which occur at indices [21, 23, 37, 73, 80, 109, 123, 127, 194, 208].

I searched each substring of the 3boat10 as the same length as the corpus, looking for an instance where characters at the listed indices were identical. The found substring was:

```
    whatisufferinthatwaynotonguecantellfrommyearliestinfancyihaveb
eenamartyrtoitasaboythediseasehardlyeverleftmeforadaytheydidnotkn
owthenthatitwasmylivermedicalsciencewasinafarlessadvancedstatetha
nnowandtheyusedtoputitdowntolazinesswhyyouskulkinglittledevilyout
heywouldsaygetupanddosomethingforyourlivingcantyounotknowingofcou
rsethatiwasillandtheydidntgivemepillstheygavemeclumpsonthesideoft
heheadandstrangeasitmayappearthoseclumpsontheheadoftencuredmefort
hetimebeingihaveknownoneclumpontheheadhavemoreeffectuponmyliveran
dmakemefeelmoreanxioustogostraightawaythenandthereanddowhatwaswan
tedtobedonewithoutfurtherlossoftimethanawholeboxofpillsdoesnowyou
knowitoftenissothosesimpleoldfashionedremediesaresometimesmoreeff
icaciousthanallthedispensarystuffw
```