CIS 3362 9/2/2022

(0) Oops - Computer @ Home → next ~~Wed~~ Fri
(showing Cryptool.java)

all files in dir
javac Cryptool.java
java Cryptool

(1) Kasiski Test
(2) Index of Coincidence
→ find keyword length

(3) Mutual Index of Coincidence

plaintext: p_0 p_1 p_2 p_3 p_4
key: k_0 k_1 k_2 k_0 k_1 k_2 ...

If we knew keyword length, we could create $|k|$ bins each which should mirror frequencies of Eng.

bin 0 c_0, c_3, c_6 ...
bin 1 c_1, c_4, c_7 ...
bin 2 c_2, c_5, c_8 ...

$|k| = 3$

Kasiski Test

Look for repeated "trigrams" or bigger.

length of keyword divides into (167-47)

$len(k) \mid 120$.

Find all repeats

w_1 : 47, 167, 267 → 120, 100

w_2 : 85, 115 → 30

w_3 : 122, 262, 402 → 140, 160

gcd(120, 100, 30, 140, 160) = 10

Index of Coincidence

Given a multiset of items

f_1 10 Skittles

f_2 5 Snickers

f_3 25 mms

f_4 10 gummyworms

50

k unique

$f_1 + f_2 + \dots f_k = n$

Choose 2 items @ random.

What's the probability they are the same?

Ans: IoC({10, 5, 25, 10})

$$IofC = \frac{10}{50} \times \frac{9}{49} + \frac{5 \times 4}{50 \times 49} + \frac{25 \times 24}{50 \times 49} + \frac{10 \times 9}{50 \times 49}$$

$$= \frac{90 + 20 + 600 + 90}{50 \times 49} = \frac{800}{50 \times 49} = \boxed{\frac{16}{49}}$$

$$IofC = \frac{\sum_{i=1}^{k} f_i(f_i - 1)}{n(n-1)}$$

random : 5 letters $\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}$ IofC ~ $\frac{1}{5}$

$\frac{1}{2}, \frac{3}{8}, \frac{1}{24}, \frac{1}{24}, \frac{1}{24}$ ~ $\frac{1}{2} \times \frac{1}{2}, \frac{3}{8} \times \frac{3}{8}$

$= \frac{1}{4} + \frac{9}{64} = \frac{25}{64}$

random 26 letters IofC $\sim \frac{1}{26}$

English $\sim .0676$

IofC is independent of which letter has which freq.

(1) Guess keyword length $|k| = 3, 4, 5, 6, \ldots$

(a) for each guess partition the cipher text into the appropriate bins.

(b) Calculate IofC of each bin

if each IofC is closer to .0676 than $\frac{1}{26} = .0378?$, then this is probably the keyword length (BREAK)

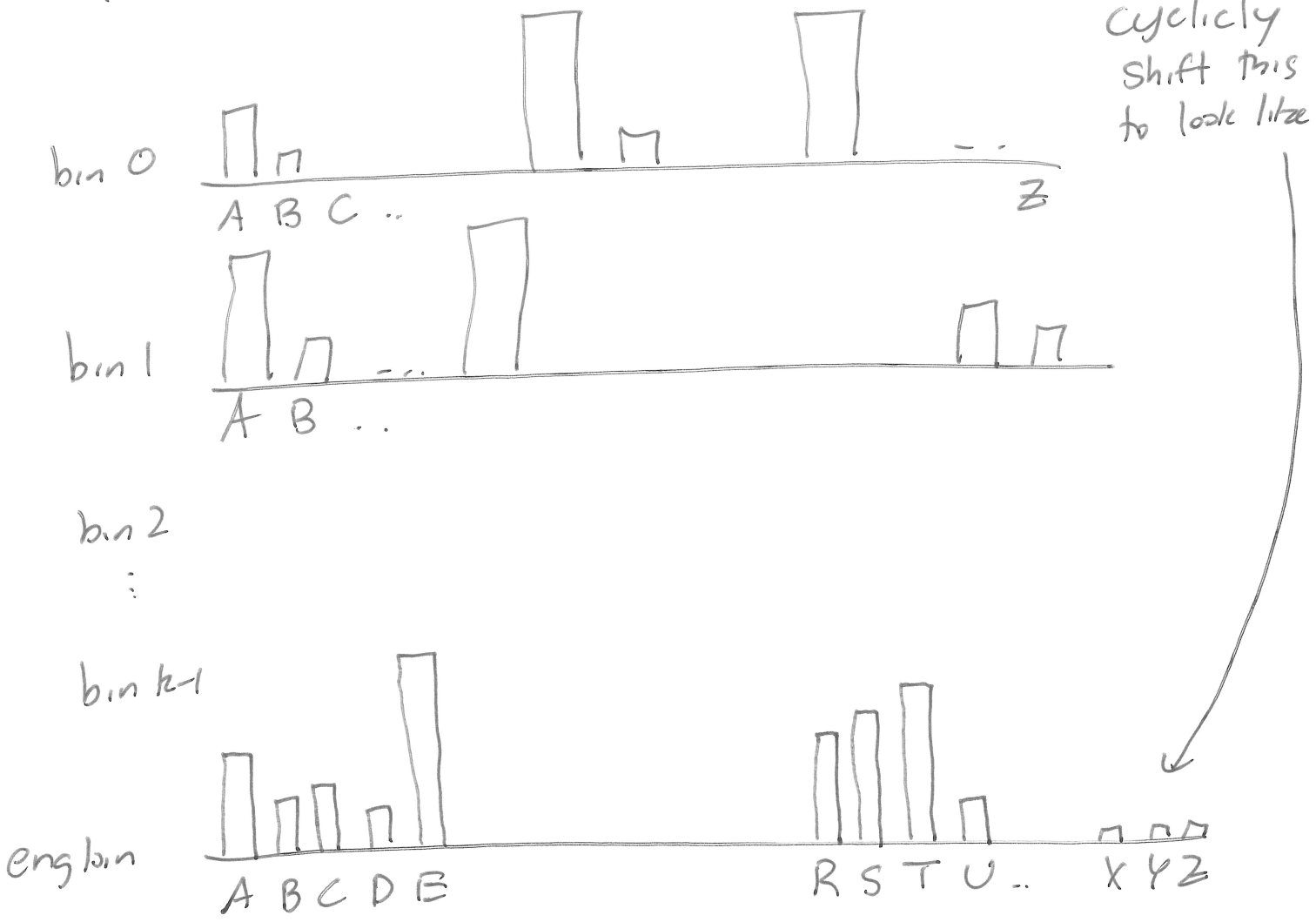
---

bins	0,	1,	2, ... $k-1$ → keyword length
A	0	2	5
B	3	0	0
C	1	0	0
D	6	0	0
⋮	⋮	⋮	⋮

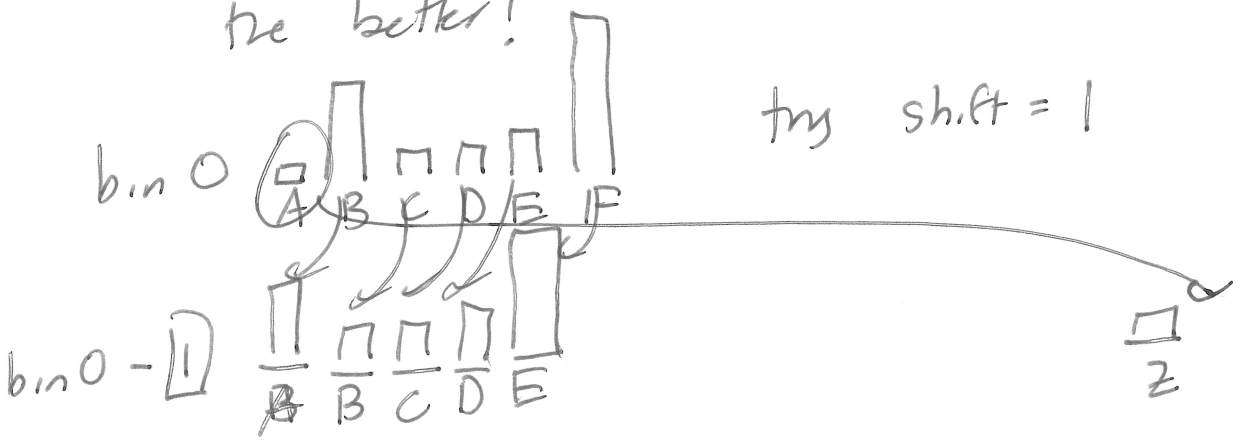
k arrays of size 26

Z

To visualize the data



Try 26 shifts.
For each shift, line up the bar graph for the bin w/ English. The "closer" they look the same the better!



Mutal Index of Coincidence

2 multisets

10 Sk	5 Sk
5 Gom	25 Gom
25 mrm	10 mrm
10 Sn	10 Sn
50	50

$$= \frac{10\times5 + 5\times25 + 25\times10 + 10\times10}{50\times50}$$

$n = f_1 + f_2 + \ldots + f_k$ Set 1

$m = g_1 + g_2 + \ldots + g_k$ set

$$MIC = \frac{\sum_{i=1}^{k} f_i \cdot g_i}{n \cdot m}$$

Of the 26 possible shifts, whichever maximizes MIC is likely to be the shift of that bin, I.E. the letter for that keyword position!