ECE5560 - Homework #2

Brian Hulette

10/6/2014

I wrote a short script, `HW2Code_Hulette_B_hulettbh.jl`, using Julia to complete this assignment. Julia is a relatively young language that is both easy to use and fast - It's quickly become very popular for scientific computing. More information can be found at `http://julialang.org/` and installers for various operating systems can be downloaded from `http://julialang.org/downloads/`

My script has two modes of operation. If you run it with one text file as an argument, like so:

```
$ julia HW2Code\_Hulette\_B\_hulettbh.jl tale_of_two_cities.txt
```

It will analyze the letter, bigram, and trigram frequency and print the results.
If you run it with two text files as arguments:

```
$ julia HW2Code\_Hulette\_B\_hulettbh.jl tale_of_two_cities.txt cipher.txt
```

It will analyze the letter, bigram, and trigram frequency of each file, then compare them to attempt to determine the key used to encrypt the second file. It uses this "first draft" key to decrypt the cipher. It will then compare the decrypted words to dictionary words to attempt to correct the key.

In this document I have compiled the output from the program into nice tables. The raw output for each case can be found in `raw_output.txt`.

b

I used `HW2Code_Hulette_B_hulettbh.jl` to analyze the frequencies in *A Tale of Two Cities* and *Moby Dick*. The output is shown in the tables below:

Letters		Bigrams		Trigrams		Letters		Bigrams		Trigrams	
e	74839	th	18166	the	11686	e	118967	th	31744	the	20029
t	54050	he	17487	and	5881	t	89548	he	26750	ing	8158
a	48168	in	11035	ing	4337	a	78956	in	20051	and	7985
o	46537	er	10493	her	3290	o	70697	an	15343	his	4094
n	42385	an	10145	his	2675	n	66669	er	15200	hat	3923
i	41017	re	8349	hat	2427	i	66586	re	11943	her	3582
h	38974	nd	7933	tha	2229	s	65011	ha	11649	tha	3575
s	37587	ha	6723	ere	2204	h	63443	nd	11116	ere	3159
r	37211	on	6717	you	2057	r	53340	at	10472	all	2965
d	28046	ou	6570	was	1799	l	43297	ng	10068	for	2783
l	22048	at	6386	for	1740	d	38769	hi	9658	ter	2612
u	16737	ed	6356	ith	1643	u	27217	on	9516	thi	2583
m	15296	en	6054	wit	1598	m	23653	ed	9356	wha	2541
w	14121	hi	5815	ent	1500	c	23121	en	9113	ver	2295
c	13899	to	5797	ion	1413	w	22500	st	8794	ent	2249
f	13563	it	5740	ver	1407	g	21239	is	8661	ith	2243
g	12547	is	5714	had	1380	f	21228	ou	8633	hal	2143
y	12185	ng	5659	thi	1317	p	17711	le	8509	wit	2099
p	9960	or	5193	ter	1303	y	17209	es	8247	but	1923
b	8422	ar	4968	not	1286	b	17165	or	8185	ale	1909
v	5204	as	4876	all	1280	v	8721	al	8164	ain	1787
k	4787	es	4626	ght	1218	k	8196	it	8050	one	1718
x	723	of	4484	him	1210	q	1567	ar	7978	ght	1707
j	714	st	4481	one	1169	j	1176	of	7458	was	1691
q	666	te	4448	our	1103	x	1062	te	7448	igh	1568
z	215	ne	4241	ess	1075	z	636	to	7360	not	1568
		se	4231	tio	1064			as	7072	ted	1533
		ve	3815	ave	1059			se	6827	eve	1529
		ea	3718	ear	1041			ea	6498	ome	1528
		me	3664	igh	1029			ll	6070	ion	1524
		le	3528	man	1010			me	5857	ate	1521
		de	3343	ted	967			wh	5819	ess	1504
		nt	3159	eve	938			ve	5815	out	1499
		ro	3097	nce	919			nt	5497	hou	1496
		ad	3092	res	905			ne	5405	tho	1491
		ti	3078	wer	893			ti	5172	hin	1440
		ll	3069	oul	888			de	5050	nce	1435
		be	2992	ome	872			ro	4960	hen	1387
		wa	2968	ain	869			ee	4714	ear	1373
		no	2945	out	865			li	4678	you	1367

c

As you can see the letter frequencies are not quite the same. If you look at the single letter order, you only have to go to the 7th letter to find a difference between the two texts. The trigrams are even worse, the second and third items, "ing" and "and", swap places between the two texts. However, we can see that certain things are pretty consistent: "e", "th" and "the" are the most common, for example.

c

If I run my program with the provided cipher text I find the following frequencies:

Letters		Bigrams		Trigrams	
x	161	jp	66	jpx	50
t	126	px	52	mtn	26
j	123	bt	43	btw	23
m	114	mt	42	hbt	15
p	107	tn	29	pbr	8
c	96	tw	26	tjx	8
r	87	xv	25	apm	7
b	83	ci	18	jpe	6
n	68	pm	18	mgg	6
v	64	cu	16	btj	6
g	49	vx	16	xvr	6
w	38	hb	15	jxv	6
y	30	br	15	pxv	6
a	29	jx	14	mlx	6
l	27	xt	13	xvx	6
i	26	pb	12	jvc	5
u	25	pc	12	pmg	5
h	23	xn	12	mrj	5
f	18	mr	12	rpm	5
e	15	gx	12	vxm	5
q	12	tj	11	pmt	5
d	6	mg	10	cut	5
o	6	gc	9	ypc	4
k	3	tc	9	cuf	4
s	1	cj	9	vbj	4
z	1	rj	9	gxj	4
		xj	9	vxj	4
		rp	9	gxn	4
		lx	9	cjp	4
		gg	8	xvq	4
		mj	8	ymr	4
		rm	8	jxt	4
		ap	8	acu	4
		ct	8	pcu	4
		cv	8	vqv	4
		jb	7	tax	4
		vc	7	qvx	4
		ml	7	twx	4
		xm	7	tcy	4
		wx	7	mta	3

I can then use the second mode of my program to find the actual key. It found:

plain letter	cipher letter
a	c
b	i
c	o
d	v
e	y
f	b
g	l
h	k
i	f
j	t
k	q
l	m
m	a
n	d
o	z
p	h
q	p
r	s
s	j
t	n
u	u
v	r
w	g
x	e
y	w
z	x

Which yields the decrypted plaintext:

in the same hour came forth fingers of a man's hand, and wrote over against the candlestick upon the plaster of the wall of the king's palace; and the king saw the part of the hand that wrote. then the king's countenance was changed, and his thoughts troubled him, so that the joints of his loins were loosed, and his knees smote one against another. the king cried aloud to bring in the astrologers, the chaldeans, and the soothsayers. and the king spake, and said to the wise men of babylon, whosoever shall read this writing, and show me the interpretation thereof, shall be clothed with scarlet, and have a chain of gold about his neck, and shall be the third ruler in the kingdom. then came in all the king's wise men; but they could not read the writing, nor make known to the king the interpretation thereof. then was king belshazzar greatly troubled, and his countenance was changed in him, and his lords were astonished. now the queen, by reason of the words of the king and his lords, came into the banquet house; and the queen spake and said, o king, live forever; let not thy thoughts trouble thee, nor let thy countenance be changed; there is a man in thy kingdom, in whom is the spirit of the holy gods; and in the days of thy wisdom of the gods, was found in him; whom the king nebuchadnezzar thy father, the king, i say, thy father, made master of the magicians, astrologers, chaldeans, and soothsayers; forasmuch as an excellent spirit, and knowledge, and understanding, interpreting of dreams, and showing of hard sentences, and dissolving of doubts, were found in the same daniel, whom the king named belteshazzar; now let daniel be called, and he will show the interpretation.