CO 480 Lecture 1

Introduction and Dr. William Thomas Tutte

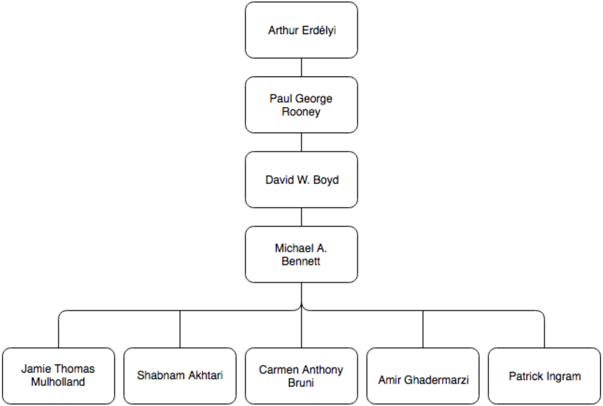
May 2nd, 2017

CO 480 - What is it

- History of mathematics is a hybrid class intersecting classical studies with mathematics in an interesting way.
- Academic history is something we all have after going to school

My Doctorate Academic Lineage

https://genealogy.math.ndsu.nodak.edu/



Arthur Erdélyi

Paul George Rooney

David W. Boyd

Michael A. Bennett

Jamie Thomas Mulholland

Shabnam Akhtari

Carmen Anthony Bruni

Amir Ghadermarzi

Patrick Ingram

Course Structure

- LEARN (for assignments, copyrighted material from other professors, etc.)
- Crowdmark (for assignment submissions and some components of the project submission)
- `http://www.cemc.uwaterloo.ca/~cbruni/CO480Resources/index.php` for lecture information (because education should be publicly available). Link is available on LEARN.
- Piazza (for course communication).

Course Structure

- The course will be broken into a series of vignettes.
- A *vignette* is a brief description of an event or episode.
- For us, our vignettes will consist of three components:
    1. A Person
    2. A Place
    3. A Problem

Course Structure

- Assignments $5\% \cdot 4 = 20\%$ (via Crowdmark)
- Two quizzes $10\% \cdot 2 = 20\%$
  Dates: Tuesday June 6th and Tuesday July 25th
- Final Project 60%

Assignments and Quizzes

- Assignments have a history component and a mathematical component
- Please use book citations as much as possible in this course.
- Quizzes will be both historical and mathematical.
- Must pass total of quizzes to pass course.

Final Project

In groups of up to 4 (and ideally a minimum of 4), you will be submitting a vignette that is to be done using LaTeX. You will be submitting the following

- Project Proposal 2% (via Crowdmark)
- Annotated Bibliography 8% (via Crowdmark)
- First Edition 25% (via LEARN Dropbox)
- Editorial Review 15% (via Crowdmark and LEARN Dropbox)
- Final Edition 10% (via LEARN Dropbox)

Project Proposal

- Pick a group (closer to size 4 is highly recommended and encouraged). Use Piazza to connect with other students.
- Pick a project - a problem, place and person born after the 1600s (these must be linked somehow).
- Consult with me if you are unsure of an idea or have a radical idea you would like to do.
- Submit your project proposal (see samples online). Due Tuesday May 16th
- To make life easier, each of you must submit these to Crowdmark (you may submit the same file of course!) - currently Crowdmark doesn't support group projects (they're working on it!)

Annotated Bibliography

- This must involve at least 5 in print sources (so you might need to go to a library!) Due Thursday June 1st.
- As before, to make life easier, each of you must submit these to Crowdmark (you may submit the same file of course!) - currently Crowdmark doesn't support group projects (they're working on it!)

First Edition

- A complete version of your 16-20 page final project (plus title and bibliography page(s)).
- Aim 'low' (realistic!) This project can become very time consuming if you are not careful. Even in groups.
- Due Tuesday June 20th.
- Submit your pdf created in LaTeX to LEARN's Dropbox - one with your names on it and one without your names.
- Sample projects can be found on LEARN as well as more detailed instructions.
- Title page and bibliography samples can be found on the CO 480 webpage.

Editorial Review

- I will randomly assign you each to edit a project.
- **Each** of you will **individually** submit an editorial review of another project. This needs to be done on your own.
- This editorial review should not contain any information about you inside it.
- Must be created in $\LaTeX$ and submitted in pdf format.
- You will then upload the Editorial review to Crowdmark and to LEARN's dropbox feature (for redistribution to the groups).
- You will be marked on completeness, accuracy, and on constructivism of your comments towards the project.
- More information can/will be found on LEARN.
- Due Thursday July 6th.

Final Edition

- Take the comments from the Editorial Review and then compose a Final Edition.
- To be submitted using LEARN's Dropbox feature.
- Due on Tuesday July 25th.

First vignette

William Thomas Tutte

- Born May 14th in 1917 at Newmarket, England
- Died May 2nd, 2002 at age 84 here in Kitchener Ontario
- Dan Younger: “The leading mathematician in combinatorics for three decades”.
- FRSC, FRS, OC [Order of Canada] 2001

More on the Life of Tutte

- Chemistry student in Cambridge in 1941
- Switched and received doctorate from Cambridge in 1948.
- Immediately after, began a teaching position at the University of Toronto. (Invited by Coxeter)
- Moved to the University of Waterloo in 1962.

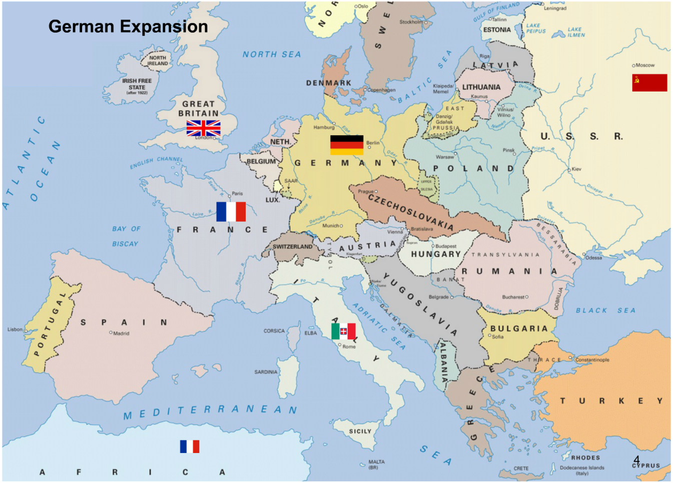
More on the Life of Tutte

- Chemistry student in Cambridge in 1941
- Switched and received doctorate from Cambridge in 1948.
- Immediately after, began a teaching position at the University of Toronto. (Invited by Coxeter)
- Moved to the University of Waterloo in 1962.

What happened between between 1941 and 1948?

Interactive map

http://www.strangecosmos.com/content/item/140489.html



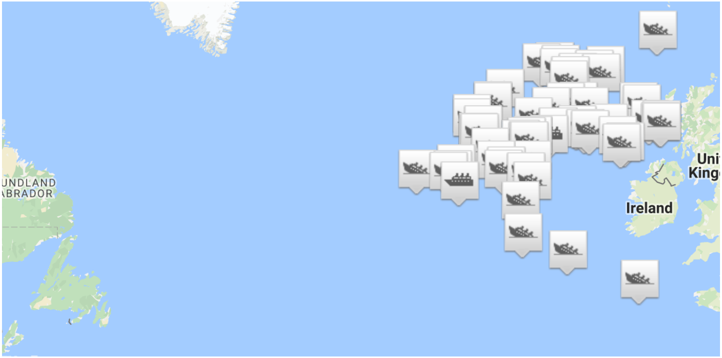
Canada's Involvement in the War

- At first, we (the Allies [of Britain]) played an auxiliary role by providing supplies to Britain.
- However, the German army began countering by attacking supply ships using U-boats

`http://www.uboataces.com/cgi-bin/uboat-photo.cgi?uniqky=2006214753182388`

Map of Losses

`http://uboat.net/allies/merchants/losses_year.html?qdate=1940-09`



What Does a Mathematician Have to do with War?

Bletchley Park (directed by Alastair Denniston):

Bletchley Park

Public domain from Wikimedia.com

Bletchley Park Mansion

Portrait photos are public domain

http://spartacus-educational.com/Alastair_Denniston.htm

Alastair Denniston

- Born in Greenock Scotland December 1st, 1881 .
- Educated at Bowdon College before studying in Bonn and Paris.
- Spoke German, an asset during the war.
- Recruited by the Admiralty. (Commanded Royal Navy in England)
- Cryptographers didn't exist but a mathematical mind was widely believed to be the best foundation.

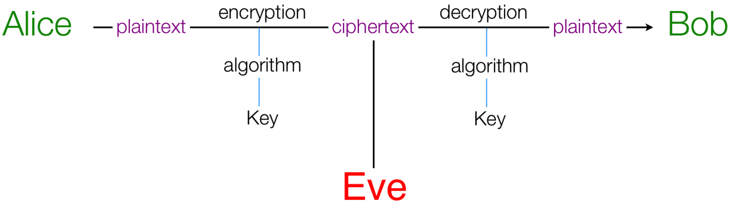
Alastair Denniston

Francis Harry Hinsley later claimed:

> *"Denniston... recruited the wartime staff from the universities with visits there in 1937 and 1938 (also 1939 when he recruited me and 20 other undergraduates within two months of the outbreak of war). I believe this was a major contribution to the wartime successes - going to the right places and choosing the right people showed great foresight."*

Francis Harry Hinsley, quoted by Robin Denniston, the author of Thirty Secret Years (2007) page 24

Cryptography



Substitution Cipher

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Q	M	W	N	B	E	R	V	T	C	Y	X	U	Z	I	L	O	K	P	H	A	G	S	F	J	D

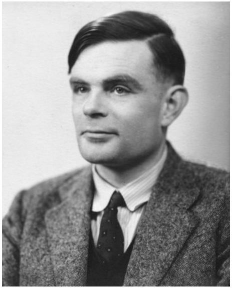
Plaintext

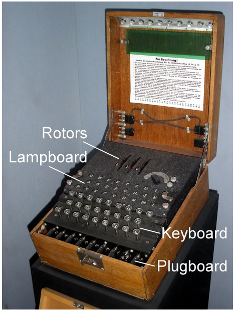
```
We shall fight on the beaches,
We shall fight on the landing grounds,
We shall fight in the fields and in the streets,
We shall fight in the hills;
We shall never surrender
```

Ciphertext

```
SB PVQXX ETRVH IZ HVB MBQWVBP
SB PVQXX ETRVH IZ HVB XQZNTZR RKIAZNP
SB PVQXX ETRVH TZ HVB ETBXNP QZN TZ HVB PHKBBHP
SB PVQXX ETRVH TZ HVB VTXXP
SB PVQXX ZBGBK PAKKBZNBK
```

Enigma and Alan Turing

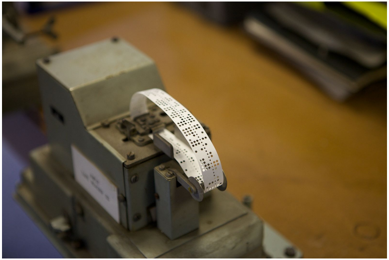




https://commons.wikimedia.org/wiki/File:EnigmaMachineLabeled.jpg

- https://www.youtube.com/watch?v=ASfAPOiq_eQ
- Enigma (2001)
- The Imitation Game (2014)

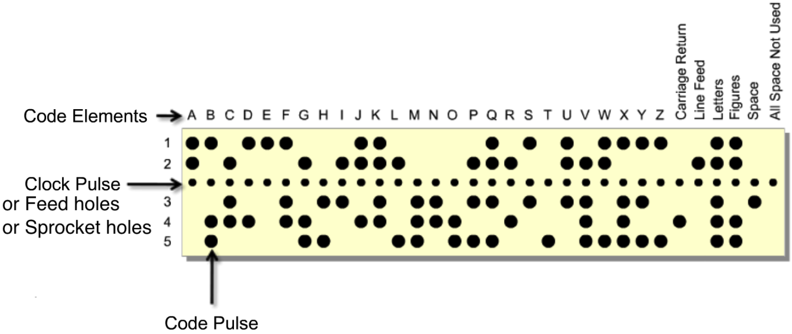
Harwell-Dekatron Computer and Teletype



Photographs of the Harwell-Dekatron Computer, also known as WITCH (Wolverhampton Instrument for Teaching Computing from Harwell). Taken at the National Museum of Computing at Bletchley Park, UK. (Wikipedia)

`https://commons.wikimedia.org/wiki/File:PaperTapes-5and8Hole.jpg`

Teletype Components



Tutte at Bletchley Park

- Responsible for decoding FISH - the teletype from German encoding
- Allies did not possess the device used to encode FISH, nor did they have any knowledge of the architecture of the machine.
- FISH carried the highest level of intelligence; encrypted communication between German High Command and Army Group headquarters in the field.
- Tutte was assigned to FISH in October 1941.

Tutte has Problems

- Inexperienced 24 year old university student
- Has no information on what the machine looks like
- Doesn’t speak German
- Doesn’t know what the message is about
- Based on radio intercepts that might introduce errors
- No computing power (other than paper and pencil!)

Even more problems!

- Axis has advantage on battlefield so there is immense pressure to break the code. USSR invaded in 1941 and suffered huge losses.
- Even if Tutte figures out what the device looks like, he still has to decode the messages.
- There is no Wikipedia or electronic library
- He cannot discuss the problem with outsiders.

Tutte's Success

- German operator error lead to Colonel John Tiltman of Bletchley Park to be able to decipher message of roughly 4000 characters.
- Three months later, there was no progress on understanding the structure of the FISH teletype
- Major Gerry Morgan, head of the research section, gives problem to Tutte. "See what you can do with this" (FISH and I, 2000)
- Using only this message and working by hand with the teletype codes (not the letters of the alphabet!) Tutte was able to identify the internal structure and behaviour of the encryption machine.

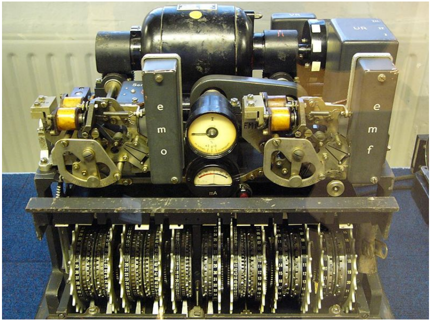
How Tutte Did this (Fish and I, 2000)

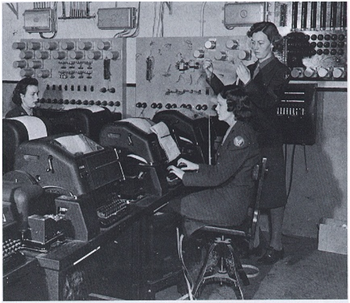
- Learnt in school that could learn something about messages by writing out period
- First 12 letters formed an “indicator” (maybe a key?)
- Knew that there “seemed to be 25 possibilities for 11 of these but only 23 for the last” (why 25 instead of 26?).
- But which to try - a period of 23? 25? How about both simultaneously! $575 = 23 \cdot 25$
- Didn’t “have much faith in this procedure but thought it best to seem busy”.

Results

- As expected, there weren't many results. However, there were lots of repeats on the diagonal!
- Tried rewriting on a period of 574 and found lots of "dot-cross patterns" of length 5 or 6. (dot=hole, cross = space)
- Tried then on a period of 41, a prime factor of 574 and got even better results. (How fortunate for Tutte that 574 was a multiple of 41 so close to 575!)
- Determined that there were two sets of wheels and wheels had different periods.

Lorenz Machine and WACs (Women's Army Corps)





WACs assigned to the Eighth Air Force in England operate teletype machines. (DOD photograph)

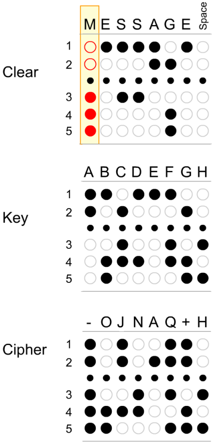
`https://en.wikipedia.org/wiki/Lorenz_cipher`
`http://www.wikiwand.com/en/Women%27s_Army_Corps`

Lorenz Machine

- 12 wheels all with a different number of teeth
- Some with irregular movements
- Arranged in three groups (Ψ, μ, χ)
- FISH had 5 bitstreams (5 hole teletype)

How FISH worked

- FISH was an additive cipher
- Uses XOR operation
- Decryption and Encryption are identical



Clear	M	E	S	S	A	G	E	Space
1	○	●	●	●	●	○	●	○
2	○	○	○	○	●	●	○	○
	•	•	•	•	•	•	•	•
3	●	○	●	●	○	○	○	○
4	●	○	○	○	○	●	○	○
5	●	○	○	○	○	●	○	○

Key	A	B	C	D	E	F	G	H
1	●	●	○	●	●	●	○	○
2	●	○	●	○	○	○	●	○
	•	•	•	•	•	•	•	•
3	○	○	●	○	○	●	○	●
4	○	●	●	●	○	●	●	○
5	○	●	○	○	○	○	●	●

Cipher	-	O	J	N	A	Q	+	H
1	●	○	●	○	○	●	●	○
2	●	○	●	○	●	●	●	○
	•	•	•	•	•	•	•	•
3	●	○	○	●	○	●	○	●
4	●	●	●	●	○	○	●	○
5	●	●	○	○	○	●	●	●

This Day in History - May 4th, 1845

William [Kingdon] Clifford born in Exeter, England. He played an important role in introducing the ideas of Riemann and other writers on non-Euclidean geometry to English mathematicians. "Clifford was a first-class gymnast, whose repertory apparently included hanging by his toes from the crossbar of a weather cock on a church tower, a feat befitting a High Churchman, as he then was."

This is the Clifford for which Clifford Algebras are named after (it is a unital associative algebra generated by a vector space and a quadratic form)
`http://www.maa.org/news/on-this-day`

Password For Online Assignment

Password for the online assignment is...

Other Announcements

- Please pick your teammates!
- Reminder: Person chosen cannot be one of the ones I'm covering in class (see the syllabus!)
- Piazza has some unmatched teammates - draft a free agent! (or maybe the free agents can team up?)
- All history essays must be typed in LaTeX.
- Project Proposal due in 12 days. (Crowdmark)
- Assignment 1 due in two weeks. (Crowdmark)
- Links for the above will come next Thursday/Friday after registration stabilizes.
- Learning Goals have been posted.
- Can't sign override forms for class size restrictions.

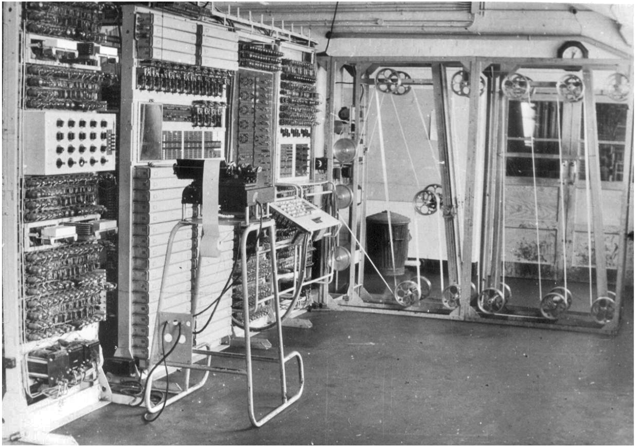
Lorenz and FISH Terminology

- New cryptographic system for longer messages for longer periods of time (unlike Enigma)
- Teletype traffic - FISH
- Lorenz machine (the machine).
- Tunny (tuna) was a word they used for the machine, the British version of the machine and even the combination of FISH and Lorenz.
- BP liked using fish words for secrets.
- `https://www.youtube.com/watch?v=b4WBINgRMTY`
- `https://www.youtube.com/watch?v=0u_9ntYRzzw`

Even With this...

- ... understanding Lorenz was one thing but still needed to find an attack on the machine!
- Tutte devised a statistical method that would work...
- ... but you needed to analyse messages thousands of times faster than one could do by hand.
- Could a machine be invented to do this?
- Tutte approached Gerry Morgan and Max Newman with the idea.

Colossus



https://commons.wikimedia.org/wiki/File:Wartime_photo_of_Colossus_10.png

Colossus

- Worked on by Max Newman and designed by Tommy Flowers of the Post Office (worked on systems for phone companies)
- Used an unprecedented number of vacuum tubes (1500)
- Existence wasn't publicly acknowledged until decades after the war (re: ENIAC)
- Prototype finished in December of 1943 operational February 5th, 1944.
- https://www.youtube.com/watch?v=9HH-asvLAj4

Colossus

- Worked on by Max Newman and designed by Tommy Flowers of the Post Office (worked on systems for phone companies)
- Used an unprecedented number of vacuum tubes (1500)
- Existence wasn't publicly acknowledged until decades after the war (re: ENIAC)
- Prototype finished in December of 1943 operational February 5th, 1944.
- https://www.youtube.com/watch?v=9HH-asvLAj4
- How important was this contribution?

Importance of Codebreaking

- Sir Harry Hinsley, official historian, British intelligence efforts in WWII:

  > *"[the war] was shortened by not less than two years and probably by four years"*

- An exhibit in 2003 on "Secret War" at the Imperial War Museum in London quoted
  British Prime Minister Winston Churchill telling King George VI

  > *"It was thanks to [codebreaking] that we won."*

(Thanks to Steve Furino for the next few slides!)

Attacking a Substitution Cipher

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z

Ciphertext (26! = 403,291,461,126,605,635,584,000,000 possible keys)

```
SB PVQXX ETRVH IZ HVB MBQWVBP
SB PVQXX ETRVH IZ HVB XQZNTZR RKIAZNP
SB PVQXX ETRVH TZ HVB ETBXNP QZN TZ HVB PHKBBHP
SB PVQXX ETRVH TZ HVB VTXXP
SB PVQXX ZBHBK PAKKBZNBK
```

Possible Plaintext

```
SB PVQXX ETRVH IZ HVB MBQWVBP
SB PVQXX ETRVH IZ HVB XQZNTZR RKIAZNP
SB PVQXX ETRVH TZ HVB ETBXNP QZN TZ HVB PHKBBHP
SB PVQXX ETRVH TZ HVB VTXXP
SB PVQXX ZBGBK PAKKBZNBK
```

Attacking a Substitution Cipher

	E						T														H				
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z

Ciphertext

```
SB PVQXX ETRVH IZ HVB MBQWVBP
SB PVQXX ETRVH IZ HVB XQZNIZG GKIAZNP
SB PVQXX ETRVH TZ HVB ETBXNP QZN TZ HVB PHKBBHP
SB PVQXX ETRVH TZ HVB VTXXP
SB PVQXX ZBHBK PAKKBZNBK
```

Possible Plaintext

```
SE PHQXX ETRHT IZ THE MEQWHEP
SE PHQXX ETRHT IZ THE XQZNTZR RKIAZNP
SE PHQXX ETRHT TZ THE ETEXNP QZN TZ THE PTKEETP
SE PHQXX ETRHT TZ THE HTXXP
SE PHQXX ZEGEK PAKKEZNEK
```

Attacking a Substitution Cipher

	E						T						D			A					H				N
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z

Ciphertext

```
SB PVQXX ETRVH IZ HVB MBQWVBP
SB PVQXX ETRVH IZ HVB XQZNIZG GKIAZNP
SB PVQXX ETRVH TZ HVB ETBXNP QZN TZ HVB PHKBBHP
SB PVQXX ETRVH TZ HVB VTXXP
SB PVQXX ZBHBK PAKKBZNBK
```

Possible Plaintext

```
SE PHAXX ETRHT IN THE MEAWHEP
SE PHAXX ETRHT IN THE XANDTNR RKIANDP
SE PHAXX ETRHT TN THE ETEXDP AND TN THE PTKEETP
SE PHAXX ETRHT TN THE HTXXP
SE PHAXX NEGEK PAKKENDEK
```

A Better Substitution Cipher

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Q	M	W	N	B	E	R	V	T	C	Y	X	U	Z	I	L	O	K	P	H	A	G	S	F	J	D

Plaintext

We shall fight on the beaches,
We shall fight on the landing grounds,
We shall fight in the fields and in the streets,
We shall fight in the hills;
We shall never surrender

Ciphertext

A Better Substitution Cipher

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Q	M	W	N	B	E	R	V	T	C	Y	X	U	Z	I	L	O	K	P	H	A	G	S	F	J	D

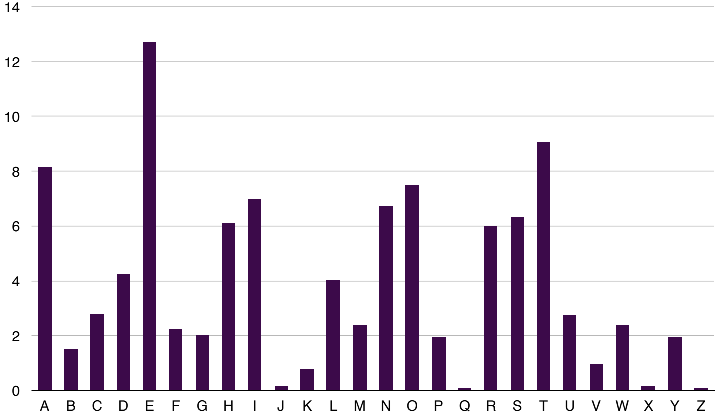
Plaintext

We shall fight on the beaches,
We shall fight on the landing grounds,
We shall fight in the fields and in the streets,
We shall fight in the hills;
We shall never surrender

Ciphertext

SBPVQ XXETR VHIZH VBMBQ WVBPS BPVQX XETRV HIZHV BXQZN TXRRK
IAZNP SBPVQ XXETR VHTZH VBETB XNPQZ NTZHV BPHKB BHPSB PVQXX
ETRVH TZHVB VTXXP SBPVQ XXZBL BKPAK KBZNB

Relatively Frequency of the English Alphabet



Letters Grouped by Relative Frequency

- E
- T, A, O, I, N, S, H, R
- D, L
- C, U, M, W, F, G, Y, P, B
- V, K, J, X, Q, Z
- Might also consider relative frequencies of letters that start or end words

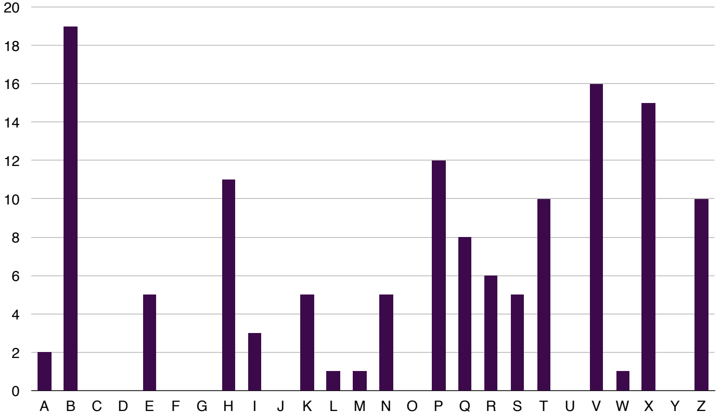
Common Pairs of Letters

- TH (3.015% of all pairs of letters)
- HE (3.004%)
- IN (1.872%)
- ER (1.860%)
- AN (1.419%)
- Might also consider pairs of identical letters: TT, LL, RR

Common Triples of Letters

- THE (2.032% of all triplets of letters)
- ING (0.747%)
- AND (0.667%)
- HER (0.547%)
- ERE (0.448%)

Relatively Frequency of the Ciphertext



Decrypting A Substitution Cipher

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
2	19	0	0	5	0	0	11	3	0	5	1	1	5	0	12	8	6	5	10	0	16	1	15	0	10

Ciphertext

```
SBPVQ XXETR VHIZH VBMBQ WVBPS BPVQX XETRV HIZHV BXQZN TXRRK
IAZNP SBPVQ XXETR VHTZH VBETB XNPQZ NTZHV BPHKB BHPSB PVQXX
ETRVH TZHVB VTXXP SBPVQ XXZBL BKPAK KBZNB
```

Plaintext

```
SBPVQ XXETR VHIZH VBMBQ WVBPS BPVQX XETRV HIZHV BXQZN TXRRK
IAZNP SBPVQ XXETR VHTZH VBETB XNPQZ NTZHV BPHKB BHPSB PVQXX
ETRVH TZHVB VTXXP SBPVQ XXZBL BKPAK KBZNB
```

Decrypting A Substitution Cipher

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
2	19	0	0	5	0	0	11	3	0	5	1	1	5	0	12	8	6	5	10	0	16	1	15	0	10
	E																								

Ciphertext

```
SBPVQ XXETR VHIZH VBMBQ WVBPS BPVQX XETRV HIZHV BXQZN TXRRK
IAZNP SBPVQ XXETR VHTZH VBETB XNPQZ NTZHV BPHKB BHPSB PVQXX
ETRVH TZHVB VTXXP SBPVQ XXZBL BKPAK KBZNB
```

Plaintext

```
SEPVQ XXETR VHIZH VEMEQ WVEPS EPVQX XETRV HIZHV EXQZN TXRRK
IAZNP SEPVQ XXETR VHTZH VEETB XNPQZ NTZHV EPHKE EHPSE PVQXX
ETRVH TZHVE VTXXP SEPVQ XXZEL EKPAK KEZNE
```

Decrypting A Substitution Cipher

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
2	19	0	0	5	0	0	11	3	0	5	1	1	5	0	12	8	6	5	10	0	16	1	15	0	10
	E																				T				

Ciphertext

```
SBPVQ XXETR VHIZH VBMBQ WVBPS BPVQX XETRV HIZHV BXQZN TXRRK
IAZNP SBPVQ XXETR VHTZH VBETB XNPQZ NTZHV BPHKB BHPSB PVQXX
ETRVH TZHVB VTXXP SBPVQ XXZBL BKPAK KBZNB
```

Plaintext

```
SEPTQ XXETR THIZH TEMEQ WTEPS EPTQX XETRT HIZHT EXQZN TXRRK
IAZNP SEPTQ XXETR THTZH TEETB XNPQZ NTZHT EPHKE EHPSE PTQXX
ETRTH TZHTE TTXXP SEPTQ XXZEL EKPAK KEZNE
```

“T” Didn’t Work, Look for “THE”

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
2	19	0	0	5	0	0	11	3	0	5	1	1	5	0	12	8	6	5	10	0	16	1	15	0	10
	E																								

Ciphertext

```
SBPVQ XXETR VHIZH VBMBQ WVBPS BPVQX XETRV HIZHV BXQZN TXRRK
IAZNP SBPVQ XXETR VHTZH VBETB XNPQZ NTZHV BPHKB BHPSB PVQXX
ETRVH TZHVB VTXXP SBPVQ XXZBL BKPAK KBZNB
```

Plaintext

```
SEPVQ XXETR VHIZH VEMEQ WVEPS EPVQX XETRV HIZHV EXQZN TXRRK
IAZNP SEPVQ XXETR VHTZH VEETB XNPQZ NTZHV EPHKE EHPSE PVQXX
ETRVH TZHVE VTXXP SEPVQ XXZEL EKPAK KEZNE
```

“T” Didn’t Work, Look for “THE”

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
2	19	0	0	5	0	0	11	3	0	5	1	1	5	0	12	8	6	5	10	0	16	1	15	0	10
	E						T														H				

Ciphertext

```
SBPVQ XXETR VHIZH VBMBQ WVBPS BPVQX XETRV HIZHV BXQZN TXRRK
IAZNP SBPVQ XXETR VHTZH VBETB XNPQZ NTZHV BPHKB BHPSB PVQXX
ETRVH TZHVB VTXXP SBPVQ XXZBL BKPAK KBZNB
```

Plaintext

```
SEPHQ XXETR HTIZT HEMEQ WHEPS EPHQX XETRH TIZTH EXQZN TXRRK
IAZNP SEPHQ XXETR HTTZT HEETB XNPQZ NTZTH EPTKE ETPSE PHQXX
ETRHT TZTHE HTXXP SEPHQ XXZEL EKPAK KEZNE
```

Realigning The Text

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
2	19	0	0	5	0	0	11	3	0	5	1	1	5	0	12	8	6	5	10	0	16	1	15	0	10
	E						T														H				

Plaintext

```
SBPVQ XXETR VHIZH VBMBQ WVBPS BPVQX XETRV HIZHV BXQZN TXRRK
IAZNP SBPVQ XXETR VHTZH VBETB XNPQZ NTZHV BPHKB BHPSB PVQXX
ETRVH TZHVB VTXXP SBPVQ XXZBL BKPAK KBZNB
```

Plaintext (Rearranged)

```
SEPHQXXETRHTIZ THE MEQWHEPSEPHQXXETRHTIZ THE XQZNTXRRKIAZNPSEPHQ
XXETRHTTZ THE ETBXNPQZNTZ THE PTKEETPSEPHQXXETRHTTZ THE
HTXXPSEPHQ XXZELEKPAKKEZNE
```

What About “XX”?

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
2	19	0	0	5	0	0	11	3	0	5	1	1	5	0	12	8	6	5	10	0	16	1	15	0	10
	E						T														H				

Plaintext

SBPVQ XXETR VHIZH VBMBQ WVBPS BPVQX XETRV HIZHV BXQZN TXRRK
IAZNP SBPVQ XXETR VHTZH VBETB XNPQZ NTZHV BPHKB BHPSB PVQXX
ETRVH TZHVB VTXXP SBPVQ XXZBL BKPAK KBZNB

Plaintext (Rearranged)

SEPHQXXETRHTIZ THE MEQWHEPSEPHQXXETRHTIZ THE XQZNTXRRKIAZNPSEPHQ
XXETRHTTZ THE ETBXNPQZNTZ THE PTKEETPSEPHQXXETRHTTZ THE
HTXXPSEPHQ XXZELEKPAKKEZNE

What About “XX”?

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
2	19	0	0	5	0	0	11	3	0	5	1	1	5	0	12	8	6	5	10	0	16	1	15	0	10
	E						T														H		L		

Plaintext

SBPVQ XXETR VHIZH VBMBQ WVBPS BPVQX XETRV HIZHV BXQZN TXRRK
IAZNP SBPVQ XXETR VHTZH VBETB XNPQZ NTZHV BPHKB BHPSB PVQXX
ETRVH TZHVB VTXXP SBPVQ XXZBL BKPAK KBZNB

Plaintext (Rearranged)

SEPHQLLETRHTIZ THE MEQWHEPSEPHQLLETRHTIZ THE LQZNTLRRKIAZNPSEPHQ
LLETRHTTZ THE ETBLNPQZNTZ THE PTKEETPSEPHQLLETRHTTZ THE
HTLLPSEPHQ LLZELEKPAKKEZNE

Vigenere Ciphers

Vigenere Ciphers

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

Vigenere Ciphers

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

As Letters

Vigenere Ciphers

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

As Letters

Plaintext	I	T		W	A	S		A		D	A	R	K		A	N	D		S	T	O
Key																					
Ciphertext																					

Vigenere Ciphers

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

As Letters

Plaintext	I	T		W	A	S		A		D	A	R	K		A	N	D		S	T	O
Key																					
Ciphertext																					

As Numbers

Vigenere Ciphers

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

As Letters

Plaintext	I	T		W	A	S		A		D	A	R	K		A	N	D		S	T	O
Key																					
Ciphertext																					

As Numbers

Plaintext	8	19		22	0	18		0		3	0	17	10		0	13	3		18	19	14
Key																					
Ciphertext																					

Vigenere Ciphers

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

As Letters

Plaintext	I	T		W	A	S		A		D	A	R	K		A	N	D		S	T	O
Key	S	N		O	O	P		Y		S	N	O	O		P	Y	S		N	O	O
Ciphertext																					

As Numbers

Plaintext	8	19		22	0	18		0		3	0	17	10		0	13	3		18	19	14
Key	18	13		14	14	15		24		18	13	14	14		15	24	18		13	14	14
Ciphertext																					

Vigenere Ciphers

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

As Letters

Plaintext	I	T		W	A	S		A		D	A	R	K		A	N	D		S	T	O
Key	S	N		O	O	P		Y		S	N	O	O		P	Y	S		N	O	O
Ciphertext	A	G		K	O	H		Y		V	N	F	Y		P	L	V		F	H	C

As Numbers

Plaintext	8	19		22	0	18		0		3	0	17	10		0	13	3		18	19	14
Key	18	13		14	14	15		24		18	13	14	14		15	24	18		13	14	14
Ciphertext	0	6		10	14	7		24		21	13	5	24		15	11	21		5	7	2

Vigenere Cipher

Plaintext

Key

Ciphertext

Vigenere Cipher

Plaintext

```
IT WAS A DARK AND STORMY NIGHT; THE RAIN FELL IN TORRENTS —
EXCEPT AT OCCASIONAL INTERVALS, WHEN IT WAS CHECKED BY A
VIOLENT GUST OF WIND WHICH SWEPT UP THE STREETS (FOR IT IS IN
LONDON THAT OUR SCENE LIES), RATTLING ALONG THE HOUSETOPS, AND
FIERCELY AGITATING THE SCANTY FLAME OF THE LAMPS THAT STRUGGLED
AGAINST THE DARKNESS.
```

Key

Ciphertext

Vigenere Cipher

Plaintext

IT WAS A DARK AND STORMY NIGHT; THE RAIN FELL IN TORRENTS — EXCEPT AT OCCASIONAL INTERVALS, WHEN IT WAS CHECKED BY A VIOLENT GUST OF WIND WHICH SWEPT UP THE STREETS (FOR IT IS IN LONDON THAT OUR SCENE LIES), RATTLING ALONG THE HOUSETOPS, AND FIERCELY AGITATING THE SCANTY FLAME OF THE LAMPS THAT STRUGGLED AGAINST THE DARKNESS.

Key SNOOPY

Ciphertext

Vigenere Cipher

Plaintext

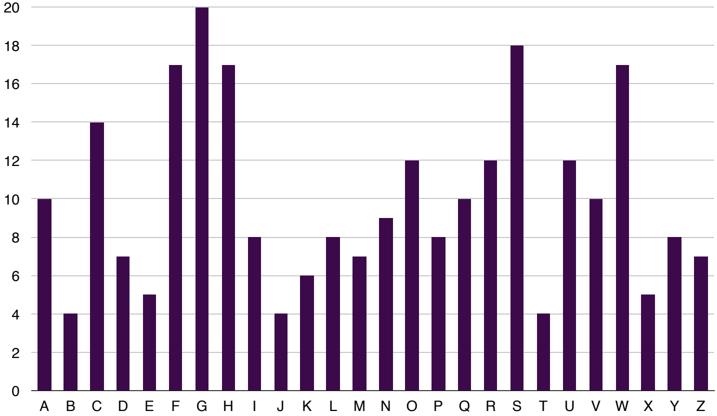
IT WAS A DARK AND STORMY NIGHT; THE RAIN FELL IN TORRENTS — EXCEPT AT OCCASIONAL INTERVALS, WHEN IT WAS CHECKED BY A VIOLENT GUST OF WIND WHICH SWEPT UP THE STREETS (FOR IT IS IN LONDON THAT OUR SCENE LIES), RATTLING ALONG THE HOUSETOPS, AND FIERCELY AGITATING THE SCANTY FLAME OF THE LAMPS THAT STRUGGLED AGAINST THE DARKNESS.

Key SNOOPY

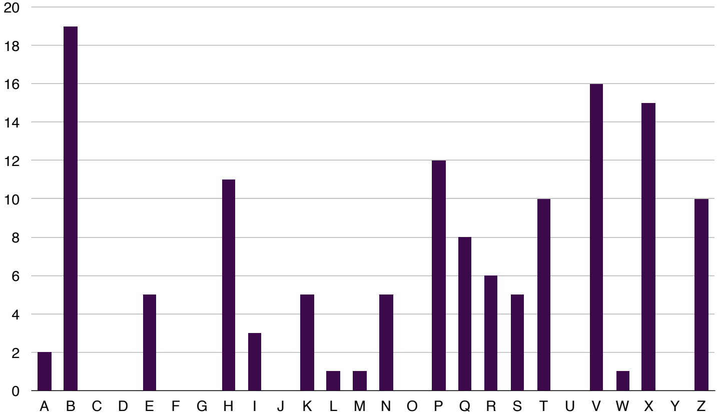
Ciphertext

```
AGKOH YVNFY PLVFH CGKQA WUWRL USFPG FSSZA GFGCF GCFGG SMAWC
HOIMU POGXM FNZWC RWEJO AQOUS BXRON GQWCU XSRQW SIWCA CFGUI
HRGSK WCBOU WQWQO RDHJN LUSGI PWRHG UMJVH WHGFY CBSMF GVOIM
MEGQT LWYWS HPSGH ZXLYN ZCCEL USVDS KRHCE QSART XCJPS ZNYYV
HOIGF THVTQ UNBHN DDNAS DDLUS ZPKHF HVPRK GFIVE DRROV YAAGH
IFWQO FZLWF G
```

Relatively Frequency of the Vigenere Ciphertext



Relatively Frequency of the Substitution Ciphertext



Substitution Cipher vs Vigenère Ciphers

- In a substitution cipher, every letter can be changed by any other letter of the alphabet so long as they are in a one to one correspondence.
- For a Vigenère cipher, each letter from 1 to the key length k is some shift of the alphabet.
- For example, the S in SNOOPY means that every sixth letter is just 18 letters further ahead in the alphabet modulo 26.

Plan for Vigenère Ciphers

- The same idea for a substitution cipher won't work since each letter gets encoded differently!
- In some sense however, these ciphertexts are really just k substitution ciphers where k is the key length.
- So if we knew the key length k, we could try to do a cryptanalysis on every kth letter, that is all the letters who's position is 1 modulo k (so $1, k+1, 2k+1, ...$). Grouping these and doing our substitution analysis is one way to attack the language.

Kasiski's Test

- Idea: If words are separated by a multiple of the key length, then they will get encoded the same way!
- Some words are common enough that they should appear multiple times in a long message.
- Let's try an example

Kasiski's Test (Example form Rosen's Elementary Number Theory)

```
QWHID  DNZEM  WTLMT  BKTIT  EMWLZ
WVCVE  HLTBS  TUDLG  WNUJE  WJEUL
EXWQO  SLNZA  NLHYQ  ALWEH  VOQWD
VQTBW  ILURY  STIJW  CLHWW  RNSIH
MNUDI  YFAVD  ELAGB  LSNZA  NSMIF
GNZEM  WALWL  CXEFA  BYJTS  SNXLH
YHULK  UCLOZ  ZAJHI  HWSM
```

Kasiski's Test (Example form Rosen's Elementary Number Theory)

```
QWHID  DNZEM  WTLMT  BKTIT  EMWLZ
WVCVE  HLTBS  TUDLG  WNUJE  WJEUL
EXWQO  SLNZA  NLHYQ  ALWEH  VOQWD
VQTBW  ILURY  STIJW  CLHWW  RNSIH
MNUDI  YFAVD  ELAGB  LSNZA  NSMIF
GNZEM  WALWL  CXEFA  BYJTS  SNXLH
YHULK  UCLOZ  ZAJHI  HWSM
```

What are some common triples of letters?

Kasiski's Test

```
QWHID   DNZEM   WTLMT   BKTIT   EMWLZ
WVCVE   HLTBS   TUDLG   WNUJE   WJEUL
EXWQO   SLNZA   NLHYQ   ALWEH   VOQWD
VQTBW   ILURY   STIJW   CLHWW   RNSIH
MNUDI   YFAVD   ELAGB   LSNZA   NSMIF
GNZEM   WALWL   CXEFA   BYJTS   SNXLH
YHULK   UCLOZ   ZAJHI   HWSM
```

Triples	Starting Positions	Differences in Starting Positions
EMW	9, 21, 129	12, 108, 120
ZEM	8, 128	120
ZAN	59, 119	60
ALW	66, 132	66
LHY	62, 149	87

So $\gcd(12, 108, 120, 60, 66, 87) = 3$ (In fact the key is USA!)

Verification

- We can also verify that the key length is correct using the *Index of Coincidence*
- Idea: Pairs of letters in English are not random since each occurs with varying frequencies.
- Probability of a random letter being a specific letter: $1/26$.
- Probability of two random letters being a specific pair of letters: $1/26^2 \approx 0.038$.

Empirically in English

- However, in English, this doesn't occur.
- Given a list of 10000 letters in a random English sentence, we expect that (approximately):

A	B	C	D	E	F	G	H	I	J	K	L	M
815	144	276	379	1311	292	199	526	635	13	42	339	254
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
710	800	198	12	683	610	1047	246	92	154	17	198	8

Index of Coincidence

- Given a random text, the number of times two letters in a text are equal is $\binom{f_i}{2} = \frac{f_i(f_i-1)}{2}$ where f_i is the frequency of letter i (for $0 \leq i \leq 25$).
- Thus, the probability that two random letters are equal (that is, that they coincide) is

$$I_C = \frac{\sum_{i=0}^{25} \binom{f_i}{2}}{\binom{n}{2}} = \frac{\sum_{i=0}^{25} f_i(f_i-1)/2}{n(n-1)/2} = \frac{\sum_{i=0}^{25} f_i(f_i-1)}{n(n-1)}$$

- The table on the previous slide can be substituted into this formula to see that $I_{\text{English}} \approx 0.065$.

Index of Coincidence

- Conclusion: If we split up the Vigenère cipher into k groupings based on the key length as before and compute the Index of Coincidence of each grouping, each grouping should have an index of coincidence "close" to that of the English language (the formula never cared that the words were valid, only that they occur in the same proportions as the English language).

References

- William T. Tutte "Fish and I" Coding Theory and Cryptography (Annapolis, MD, 1998), Springer, Berlin, 2000.
- `http://people.math.binghamton.edu/zaslav/Oldcourses/581.S03/tutte.html`
- CO 680 on LEARN videos with Dan Younger and Steve Furino
- CO 480 slides from Steve Furino
- `http://spartacus-educational.com/Alastair_Denniston.htm` (See the references there as well)