# Lecture 2 – A Look Back

CS-466 Applied Cryptography
Adam O’Neill

# A Taste of Notation

you can think of algorithms as Turing Machines

* randomized : has a random coin tape (or can make coin flips

$\underbrace{y \xleftarrow{\$} A(\ldots)}_{\text{run } A \text{ with fresh randomness}}$ $\underbrace{y \leftarrow A(\ldots, r)}_{\text{run } A \text{ with coins fixed to } r}$

* oracles :

$y \xleftarrow{\$} A^{\mathcal{O}_1(\cdot), \ldots, \mathcal{O}_n(\cdot)}(\ldots)$ } run $A$ with access to oracles $\mathcal{O}_1, \mathcal{O}_2, \ldots, \mathcal{O}_n$

If $S$ is a finite set then $|S|$ is cardinality, $s \xleftarrow{\$} S$ means $s$ is assigned a random element from $S$

$\{0,1\}^*$: set of all strings $\{0,1\}^n$ set of strings of length $n$

$|s|$ length of string $s$.

If $s_1, s_2$ are strings then $s_1 \| s_2$ denotes an encoding from which $s_1, s_2$ are uniquely recoverable

if lengths of $s_1, s_2$ are known then concatenation suffices

# Probabilities

Games (will formalize a syntax later)

Ex. Game 1

$a \xleftarrow{\$} A(\_)$

$b \xleftarrow{\$} B(a)$

$c \leftarrow a \oplus b$

$\Pr[c = x]$ means the probability that $c = x$ over the choice of randomness for the entire game

sample space $\Omega$

## Random Variables

If $X$ is a random variable it is over some finite set $\mathcal{X}$ (the possible ou~~tcomes~~s of $X$. It puts a probability distribution on $\mathcal{X}$.

$\Pr_X[\ldots]$ } $\Pr \ldots$ is true when $X$ is sampled randomly according to its distribution

# Defining Cryptographic Objects

Think of cryptographic objects the same way you think of a data type in programming

To define a cryptographic object you need to specify: e.g. encryption scheme, hash function

- Syntax: What are the constituent algorithms, what are their inputs and outputs
- Correctness: what functionality these algorithms need to fulfill in the absense of an adversary. to be useful
- Security:
  - what an adversary is allowed to do in attacking a scheme
  - what is the adversary's goal to break the scheme

typically specified by a game

# Symmetric-Key Encryption: Syntax

A symmetric-key encryption scheme with message-space $\mathcal{M}$ is a tuple of algorithms $\mathsf{SE} = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ defined as follows:

w/ message-space $\mathcal{M}$, ciphertext-space $\mathcal{C}$, coin-space $\mathcal{R}$

set of possible message

possible ciphertexts

possible coin sequences of encryption ALG

- $\mathcal{K}$ outputs a key $K$

  $K \xleftarrow{\$} \mathcal{K}$

- $\mathcal{E}$ on input a key $K$ and a message $m$ outputs a ciphertext $c$ CAN BE RANDOMIZED

  $c \xleftarrow{\$} \mathcal{E}(K, m)$ also written as $c \xleftarrow{\$} \mathcal{E}_K(m)$

- $\mathcal{D}$ on input a key $K$ and ciphertext $c$ outputs a message $m$ or special "error symbol" $\perp$ "bot" "bottom"

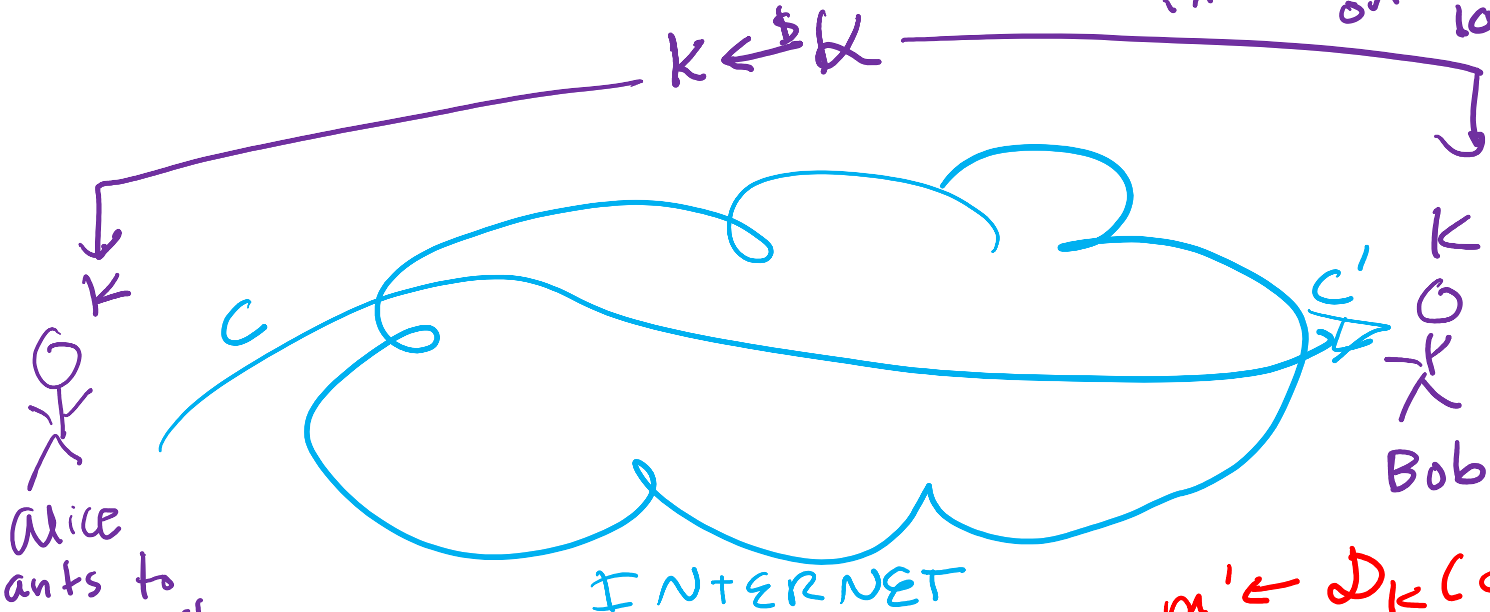
  means ciphertext is invalid

  DETERMINISTIC why??

  $m \leftarrow \mathcal{D}(K, c)$ or $m \leftarrow \mathcal{D}_K(c)$

# Usage

we will explain how parties can initially agree on $K$ later!




Alice wants to transmit a message $m$ to Bob

$c \xleftarrow{\$} \mathcal{E}_K(m)$

$m' \leftarrow \mathcal{D}_K(c')$

Does $m' = m$?

# Correctness

Fix a message m in the message space. Fix a key K that can be output by $\mathcal{K}$

Consider the game

Sample space: Coin space of $\mathcal{E}$

$c \xleftarrow{\$} \mathcal{E}_K(m)$

$m' \leftarrow \mathcal{D}_K(c)$

or "perfectly correct"

We say SE is correct if $\forall m \, \forall K$

$\Pr[m = m'] = 1$

Bob always! gets the message alice sent when nobody tampers with the ciphertext

doesn't really matter if Eve is passively observing

What are some equivalent ways of writing the correctness condition?

Weakest adversarial goal: Given some $C \leftarrow \mathcal{E}_K(M)$, recover $M$.

If this is easy the encryption scheme is clearly insecure

Adversary knows the scheme's message space $\mathcal{M}$ and algorithms $\mathcal{K}, \mathcal{E}, \mathcal{D}$ but not the key $K$. (Kerckhoff's principle)

IMPORTANT

why?

usually the adversary will figure these out anyway

If security relies on secrecy of the algorithms and there is a compromise, it is much harder to replace the entire scheme than choose a new random key K.

# Substitution Ciphers

Today's lecture concerns substitution ciphers, a particular type of symmetric-key encryption.

Prior to modern cryptography, basically all encryption schemes were of this form.

Yet we will see that such schemes are fundamentally flawed.

# Scheme Setup

Definition

could be $\Sigma^*$ itself

$$\Sigma = \{A, B, \ldots, Z\} \cup \{\sqcup, ., ?, !, \ldots\}$$

the message space must be a subset of $\Sigma^*$

the ciphertext space can be any symbols

The key generation algorithm must output a permutation on $\Sigma$

one-to-one mapping

The encryption algorithm applies $\pi$ "component-wise" to the characters in the input message, i.e. has the form (can be written as) in the following slide

To each character individually one-by-one

# Definition of S.C. Algorithms

Algorithm $\mathcal{E}_\pi(M)$
  For $i = 1, \ldots, |M|$ do
    $C[i] \leftarrow \pi(M[i])$
Return $C$

Algorithm $\mathcal{D}_\pi(C)$
  For $i = 1, \ldots, |C|$ do
    $M[i] \leftarrow \pi^{-1}(C[i])$
Return $M$

Excercise

prove correctness

# Cryptanalysis

Suppose adversary has ciphertext:

```
COXBX TBX CVK CDGXR DI T GTI'R ADHX VOXI OX ROKQAU
IKC RNXPQATCX: VOXI OX PTI'C THHKBU DC, TIU VOXI OX
PTI.
```

# Frequency

```
   E E   E        E          '     E   E   E
COXBX TBX CVK CDGXR DI T GTI'R ADHX VOXI OX ROKQAU

     E     E:   E   E    '          ,      E
IKC RNXPQATCX: VOXI OX PTI'C THHKBU DC, TIU VOXI

 E    .
OX PTI.
```

| A | B | C | D | E | F | G | H | I | J | K | L | M |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 | 3 | 7 | 4 | 0 | 0 | 2 | 3 | 9 | 0 | 4 | 0 | 0 |

| N | O | P | Q | R | S | T | U | V | W | X | Y | Z |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 8 | 3 | 2 | 4 | 0 | 8 | 3 | 4 | 0 | 13 | 0 | 0 |

```
  E E   E        E          '     E   E   E
COXBX TBX CVK CDGXR DI T GTI'R ADHX VOXI OX ROKQAU

      E     E:   E   E    '           ,       E
IKC RNXPQATCX: VOXI OX PTI'C THHKBU DC, TIU VOXI
                              L?
 E    .                       T?
OX PTI.
```

**Any ideas?**

OX in ciphertext $\Rightarrow \pi^{-1}(\texttt{O}) \in \{\texttt{B},\texttt{H},\texttt{M},\texttt{W}\}$

Guess $\pi^{-1}(\texttt{O}) = \texttt{H}$ since O has pretty high frequency

```
THE E   E T   T  E          '     E  HE  HE  H
COXBX TBX CVK CDGXR DI T GTI'R ADHX VOXI OX ROKQAU

  T   E    TE:  HE  HE    'T         T,      HE
IKC RNXPQATCX: VOXI OX PTI'C THHKBU DC, TIU VOXI

HE    .
OX PTI.
```

```
*HE*E
COXBX
```

Could be: `THERE,THESE,WHERE,...`

Guess $\pi^{-1}(C) = T$ since there is no ? in ciphertext so `WHERE` is unlikely.

```
THERE ARE T     T  E       A   A   ’        E  HE   HE  H
COXBX TBX CVK CDGXR DI T GTI’R ADHX VOXI OX ROKQAU

  T   E    ATE:  HE  HE  A ’T A    R    T, A     HE
IKC RNXPQATCX: VOXI OX PTI’C THHKBU DC, TIU VOXI

HE  A .
OX PTI.
```

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\pi^{-1}(\tau)$ | | R | T | | | | | | | | | |
| $\tau$ | N | O | P | Q | R | S | T | U | V | W | X | Y |

T is a single-letter word so $\pi^{-1}(T) \in \{\mathtt{A}, \mathtt{I}\}$
We know $\pi^{-1}(\mathtt{B}) \in \{\mathtt{R}, \mathtt{S}\}$
So TBX could be: ARE, ASE, IRE, ISE
We guess ARE

See also: https://en.wikipedia.org/wiki/Cryptanalysis_of_the_Enigma

THERE ARE T     T  E       A   A   ’        E  HE   HE

# Result

```
THERE ARE TWO TIMES IN A MAN'S LIFE WHEN HE SHOULD
COXBX TBX CVK CDGXR DI T GTI'R ADHX VOXI OX ROKQAU

NOT SPECULATE: WHEN HE CAN'T AFFORD IT, AND WHEN
IKC RNXPQATCX: VOXI OX PTI'C THHKBU DC, TIU VOXI

HE CAN.
OX PTI.
```

# Another Example

**Example 1** (Cryptography). Stanford's Statistics Department has a drop-in consulting service. One day, a psychologist from the state prison system showed up with a collection of coded messages. Figure 1 shows part of a typical example.

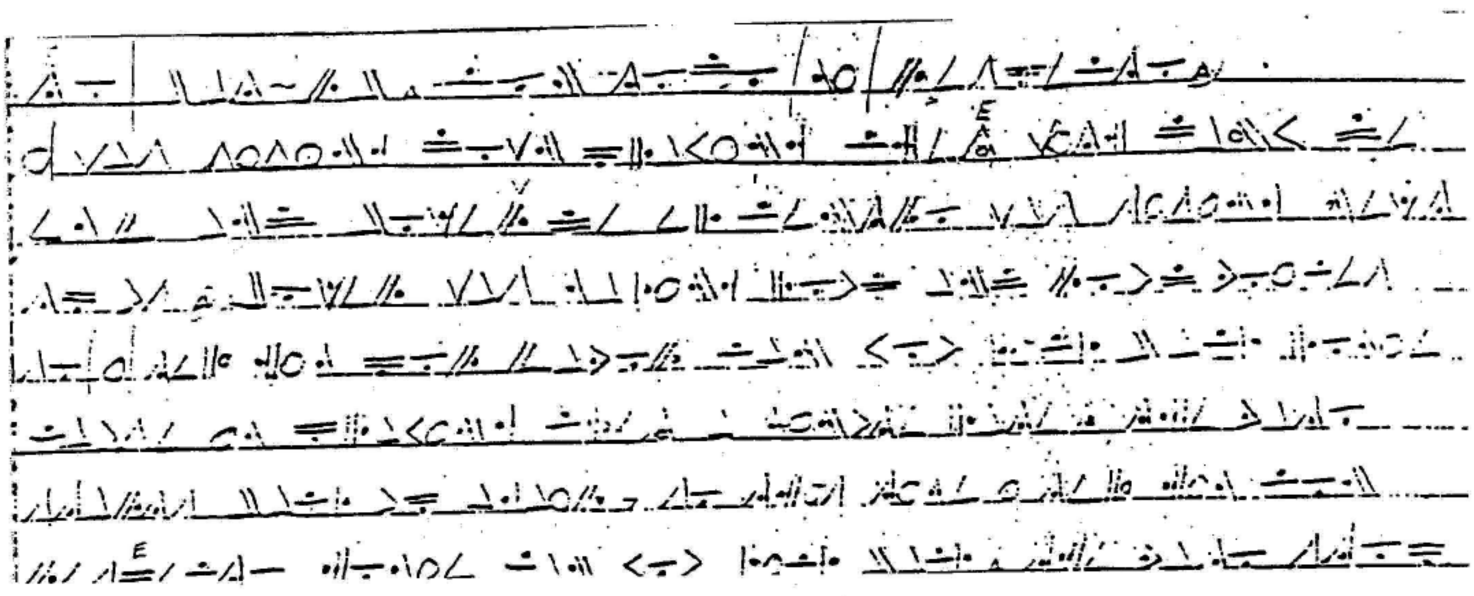


**Figure 1**

https://math.uchicago.edu/~shmuel/Network-course-readings/MCMCRev.pdf

# The General Setting

$$f : \{\text{code space}\} \longrightarrow \{\text{usual alphabet}\}.$$

To get the statistics, Marc downloaded a standard text (e.g., *War and Peace*) and recorded the first-order transitions: the proportion of consecutive text symbols from $x$ to $y$. This gives a matrix $M(x, y)$ of transitions. One may then associate a plausibility to $f$ via

i-th character in ciphertext

$$\mathrm{Pl}(f) = \prod_i M\left(f(s_i), f(s_{i+1})\right)$$

M[x,y]

likelihood (according to the training set) that x follows y

over the characters position in the ciphertext

can generalize to n-th order transitions

i.e. "n-grams"

BREAKS

"souped up" substitution ciphers like Enigma

# The Metropolis Algorithm

- Start with a preliminary guess, say $\boxed{f}$.
- Compute $\mathrm{Pl}(f)$.
- Change to $f_*$ by making a random transposition of the values $f$ assigns to two symbols.
- Compute $\mathrm{Pl}(f_*)$; if this is larger than $\mathrm{Pl}(f)$, accept $f_*$.
- If not, flip a $\mathrm{Pl}(f_*)/\mathrm{Pl}(f)$ coin; if it comes up heads, accept $f_*$.
- If the coin toss comes up tails, stay at $f$.

$f$ : symbols → usual alphabet ~~usual alph,~~

$f(n) = a$

$f(r) = c$

~~d~~on't want to get stuck in local maximum

# Result

to bat-rb. con todo mi respeto. i was sitting down playing chess with danny de emf and boxer de el centro was sitting next to us. boxer was making loud and loud voices so i tell him por favor can you kick back homie cause im playing chess a minute later the vato starts back up again so this time i tell him con respecto homie can you kick back. the vato stop for a minute and he starts up again so i tell him check this out shut the f**k up cause im tired of your voice and if you got a problem with it we can go to celda and handle it. i really felt disrespected thats why i told him. anyways after i tell him that the next thing I know that vato slashes me and leaves. dy the time i figure im hit i try to get away but the c.o. is walking in my direction and he gets me right dy a celda. so i go to the hole. when im in the hole my home boys hit doxer so now "b" is also in the hole. while im in the hole im getting schoold wrong and

# Watch the Algorithm Work

```
 100 ER ENOHDLAE OHDLO UOZEOUNORU O UOZEO HD OITO HEOQSET IUROFHE HENO ITORUZAEN
 200 ES ELOHRNDE OHRNO UOVEOULOSU O UOVEO HR OITO HEOQAET IUSOPHE HELO ITOSUVDEL
 300 ES ELOHANDE OHANO UOVEOULOSU O UOVEO HA OITO HEOQRET IUSOFHE HELO ITOSUVDEL
 400 ES ELOHINME OHINO UOVEOULOSU O UOVEO HI OATO HEOQRET AUSOWHE HELO ATOSUVMEL
 500 ES ELOHINME OHINO UODEOULOSU O UODEO HI OATO HEOQRET AUSOWHE HELO ATOSUDMEL
 600 ES ELOHINME OHINO UODEOULOSU O UODEO HI OATO HEOQRET AUSOWHE HELO ATOSUDMEL
 900 ES ELOHANME OHANO UODEOULOSU O UODEO HA OITO HEOQRET IUSOWHE HELO ITOSUDMEL
1000 IS ILOHANMI OHANO RODIORLOSR O RODIO HA OETO HIOQUIT ERSOWHI HILO ETOSRDMIL
1100 ISTILOHANMITOHANOT ODIO LOS TOT ODIOTHATOEROTHIOQUIRTE SOWHITHILOTEROS DMIL
1200 ISTILOHANMITOHANOT ODIO LOS TOT ODIOTHATOEROTHIOQUIRTE SOWHITHILOTEROS DMIL
1300 ISTILOHARMITOHAROT ODIO LOS TOT ODIOTHATOENOTHIOQUINTE SOWHITHILOTENOS DMIL
1400 ISTILOHAMRITOHAMOT OFIO LOS TOT OFIOTHATOENOTHIOQUINTE SOWHITHILOTENOS FRIL
1600 ESTEL HAMRET HAM TO CE OL SOT TO CE THAT IN THE QUENTIOS WHETHEL TIN SOCREL
1700 ESTEL HAMRET HAM TO BE OL SOT TO BE THAT IN THE QUENTIOS WHETHEL TIN SOBREL
1800 ESTER HAMLET HAM TO BE OR SOT TO BE THAT IN THE QUENTIOS WHETHER TIN SOBLER
1900 ENTER HAMLET HAM TO BE OR NOT TO BE THAT IS THE QUESTION WHETHER TIS NOBLER
2000 ENTER HAMLET HAM TO BE OR NOT TO BE THAT IS THE QUESTION WHETHER TIS NOBLER
```

# Take Away?

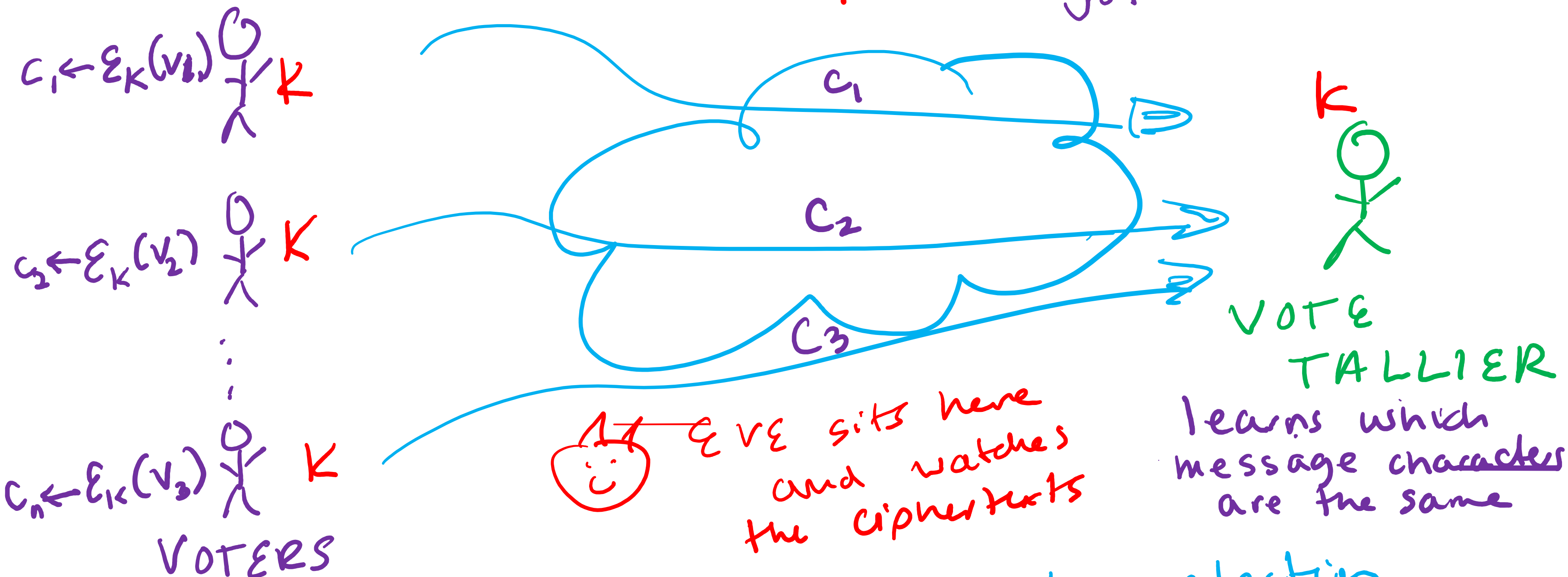
One might conclude substitution ciphers are flawed, but only for long enough plaintexts.

By thinking more adversarially, we can come up with an example that shows that such ciphers are flawed even with 1-bit plaintexts.

# Voting

Everyone shares a key $K$

voters encrypt 0/1 for "no" or "yes"
~~voter~~ $i$ has vote $v_i$

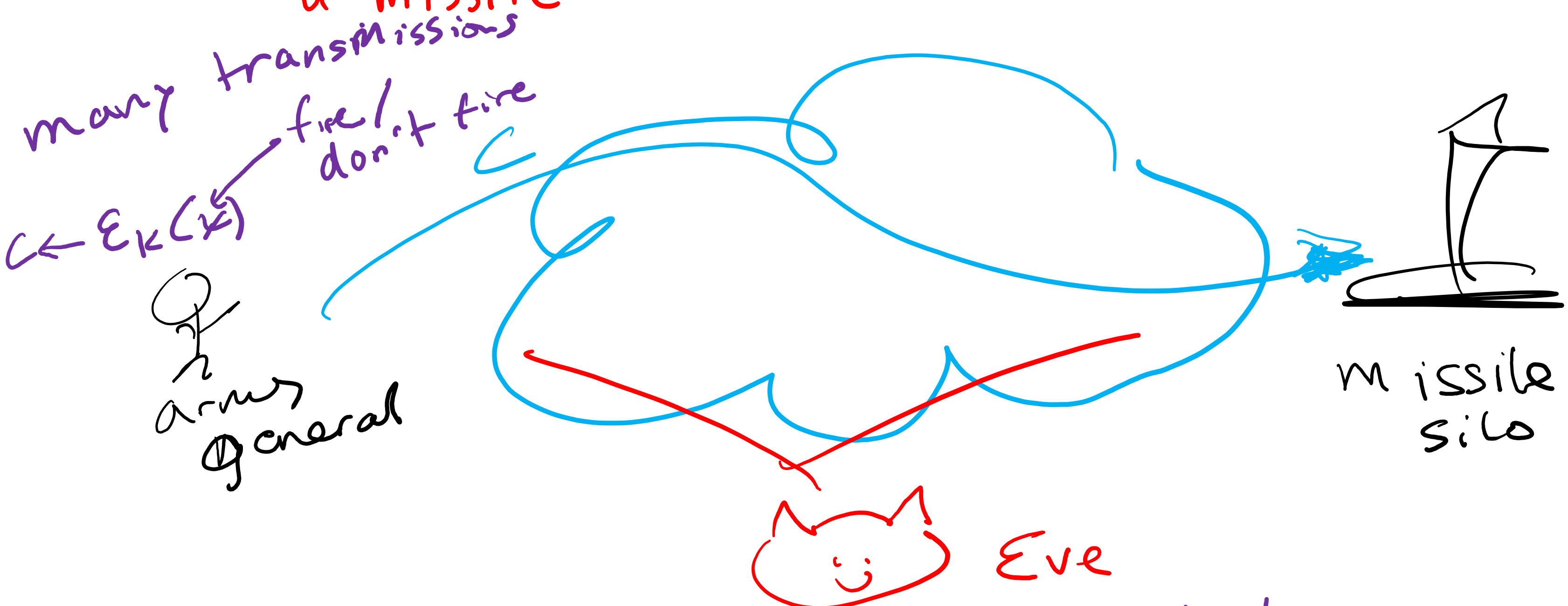



Claim: If Eve learns the election outcome then she can decrypt all ciphertexts!

If Eve is one of the voters she knows who voted the same as her (even for many choices of votes)

# Military

Suppose ~~an~~ there is a missle silo and a general transmits fire / don't fire a missile




If after the first ciphertext Eve sees what happens to the missile then she knows what will happen next time she sees a ciphertext