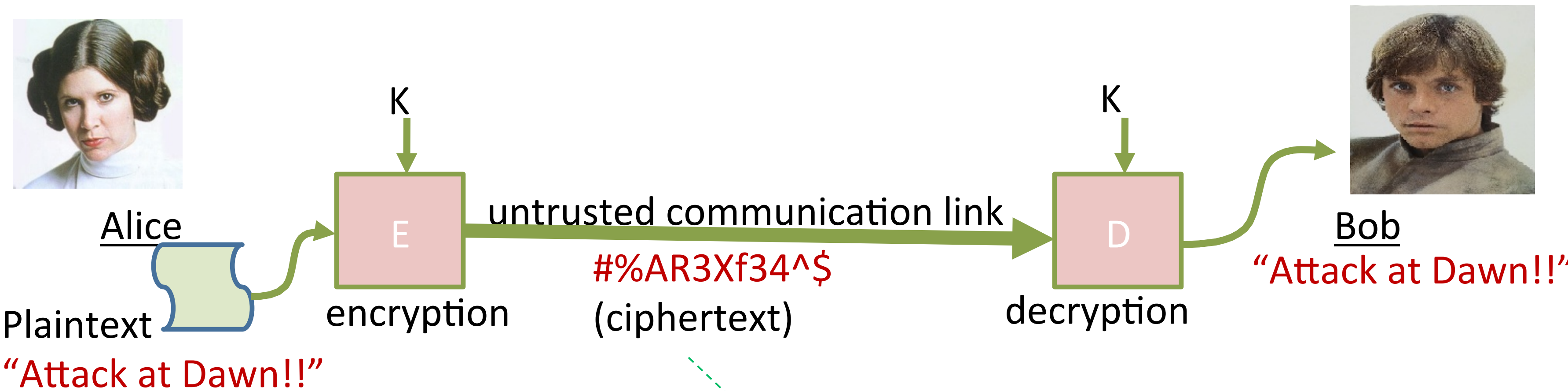
Perfect Secrecy

Chester Rebeiro
IIT Madras

Encryption



How do we design ciphers?

Cipher Models
(What are the goals of the design?)

Computation Security

My cipher can withstand all attacks with complexity less than 2^{2048}

The best attacker with the best computation resources would
take 3 centuries to attack my cipher

Provable Security (Hardness relative to a tough problem)

If my cipher can be broken then large numbers can be factored easily

Unconditional Security

My cipher is secure against all attacks irrespective of the
attacker's power.
I can prove this!!

This model is also known as **Perfect Secrecy.**
Can such a cryptosystem be built?
We shall investigate this.

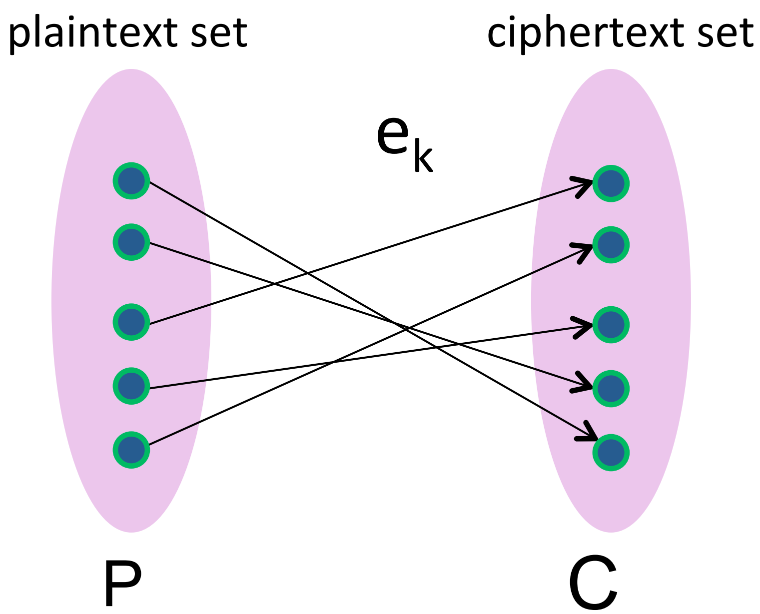
Analyzing Unconditional Security

- Assumptions
  - Ciphertext only attack model

    The attacker only has information about the ciphertext. The key and plaintext are secret.

- We first analyze a single encryption then relax this assumption by analyzing multiple encryptions with the same key

Encryption



- For a given key, the encryption (e_k) defines an injective mapping between the plaintext set (P) and ciphertext set (C)

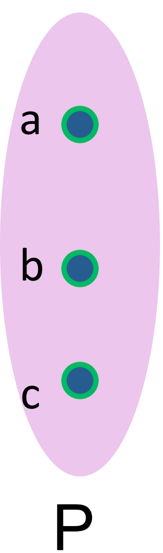
- Alice picks a plaintext $x \in P$, chooses a key (independently), and encrypts it to obtain a ciphertext $y \in C$

CR

Plaintext Distribution

Plaintext Distribution

- Let **X** be a discrete random variable over the set P
- Alice chooses x from P based on some probability distribution
  - Let *Pr[**X** = x]* be the probability that x is chosen
  - This probability may depend on the language



Plaintext set
Pr[**X**=a] = 1/2
Pr[**X**=b] = 1/3
Pr[**X**=c] = 1/6

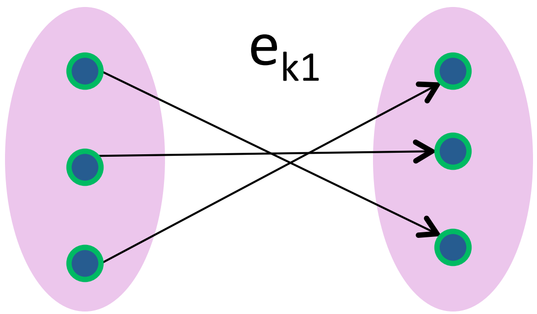
Note : Pr[a] + Pr[b] + Pr[c] = 1

Key Distribution

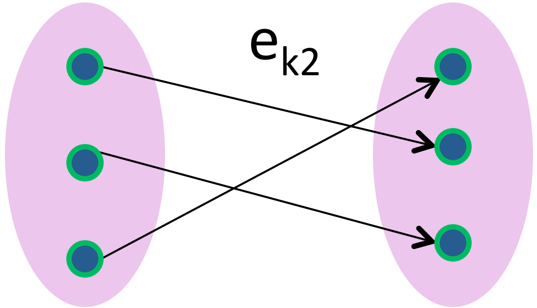
Key Distribution

- Alice & Bob agree upon a key k chosen from a key set K
- Let ***K*** be a random variable denoting this choice

keyspace
Pr[**K**=k_1] = ¾
Pr[**K**=k_2] = ¼



There are two keys in the keyset thus there are two possible encryption mappings



Ciphertext Distribution

- Let **Y** be a discrete random variable over the set C
- The probability of obtaining a particular ciphertext y depends on the plaintext and key probabilities

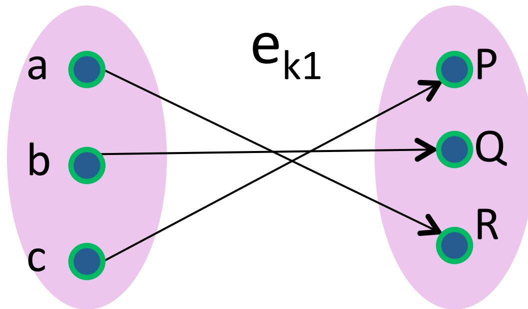
$$\Pr[Y = y] = \sum_k \Pr(k)\Pr(d_k(y))$$

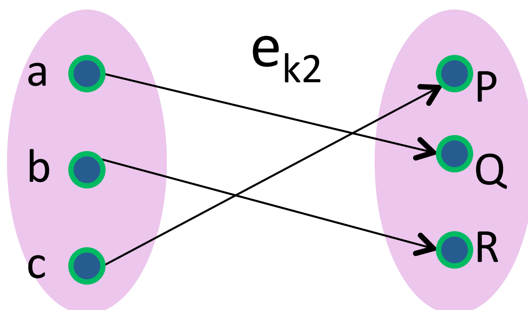
Pr[Y = P] = Pr(k_1) * Pr(c) + Pr(k_2) * Pr(c)
= (3/4 * 1/6) + (1/4 * 1/6) = **1/6**

Pr[Y = Q] = Pr(k_1) * Pr(b) + Pr(k_2) * Pr(a)
= (3/4 * 1/3) + (1/4 * 1/2) = **3/8**

Pr[Y = R] = Pr(k_1) * Pr(a) + Pr(k_2) * Pr(b)
= (3/4 * 1/2) + (1/4 * 1/3) = **11/24**

Note: Pr[Y=P] + Pr[Y=Q] + Pr[Y=R] = 1





plaintext
Pr[**X**=a] = 1/2
Pr[**X**=b] = 1/3
Pr[**X**=c] = 1/6

keyspace
Pr[**K**=k_1] = ¾
Pr[**K**=k_2] = ¼

Attacker's Probabilities

- The attacker wants to determine the plaintext ***x***
- Two scenarios
  - Attacker does not have y (a priori Probability)
    - Probability of determining ***x*** is simply ***Pr[x]***
    - Depends on plaintext distribution (eg. Language charcteristics)
  - Attacker has y (a posteriori probability)
    - Probability of determining x is simply ***Pr[x|y]***

A posteriori Probabilities

- How to compute the attacker's a posteriori probabilities? $\Pr[X = x \mid Y = y]$
  - Bayes' Theorem

$$\boxed{\Pr[x \mid y] = \frac{\Pr[x] \times \Pr[y \mid x]}{\Pr[y]}}$$

probability of the plaintext

probability of this ciphertext

?

The probability that y is obtained given x depends on the keys which provide such a mapping

$$\Pr[y \mid x] = \sum_{\{k \,:\, d_k(y)=x\}} \Pr[k]$$

Pr[y|x]

$\Pr[P|a] = 0$

$\Pr[P|b] = 0$

$\Pr[P|c] = 1$

$\Pr[Q|a] = \Pr[k_2] = ¼$

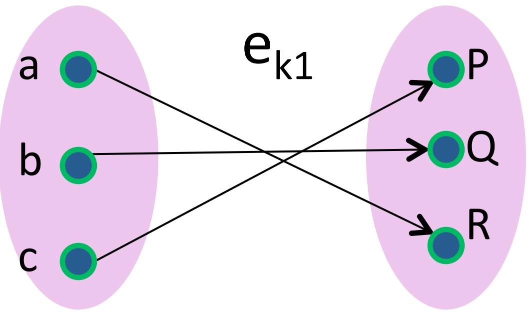
$\Pr[Q|b] = \Pr[k_1]= ¾$

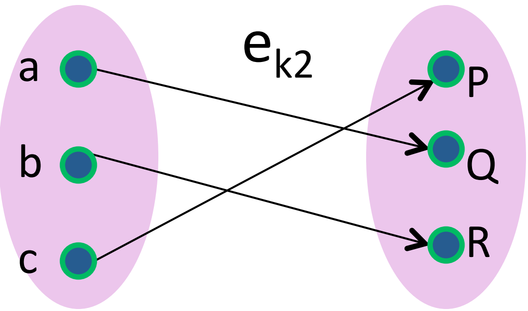
$\Pr[Q|c] = 0$

$\Pr[R|a] = \Pr[k_1] = ¾$

$\Pr[R|b] = \Pr[k_2] = ¼$

$\Pr[R|c] = 0$





keyspace
$\Pr[\mathbf{K}=k_1] = ¾$
$\Pr[\mathbf{K}=k_2] = ¼$

Computing A Posteriori Probabilities

$$\Pr[x \mid y] = \frac{\Pr[x] \times \Pr[y \mid x]}{\Pr[y]}$$

plaintext	ciphertext
Pr[**X**=a] = 1/2	Pr[**Y**=P] = 1/6
Pr[**X**=b] = 1/3	Pr[**Y**=Q] = 3/8
Pr[**X**=c] = 1/6	Pr[**Y**=R] = 11/24

Pr[y|x]
Pr[P|a] = 0 Pr[P|b] = 0 Pr[P|c] = 1
Pr[Q|a] = ¼ Pr[Q|b] = ¾ Pr[Q|c] = 0
Pr[R|a] = ¾ Pr[R|b] = ¼ Pr[R|c] = 0

Pr[a|P] = 0 Pr[b|P] = 0 Pr[c|P] = 1

Pr[a|Q] = 1/3 Pr[b|Q] = 2/3 Pr[c|Q] = 0

Pr[a|R] = 9/11 Pr[b|R] = 2/11 Pr[c|R] = 0

If the attacker sees ciphertext ***P*** then she would know the plaintext was ***c***

If the attacker sees ciphertext ***R*** then she would know ***a*** is the most likely plaintext

Not a good encryption mechanism!!

Perfect Secrecy

- Perfect secrecy achieved when

a posteriori probabilities = a priori probabilities

$$\boxed{\Pr[x \mid y] = \Pr[x]}$$

i.e the attacker learns nothing from the ciphertext

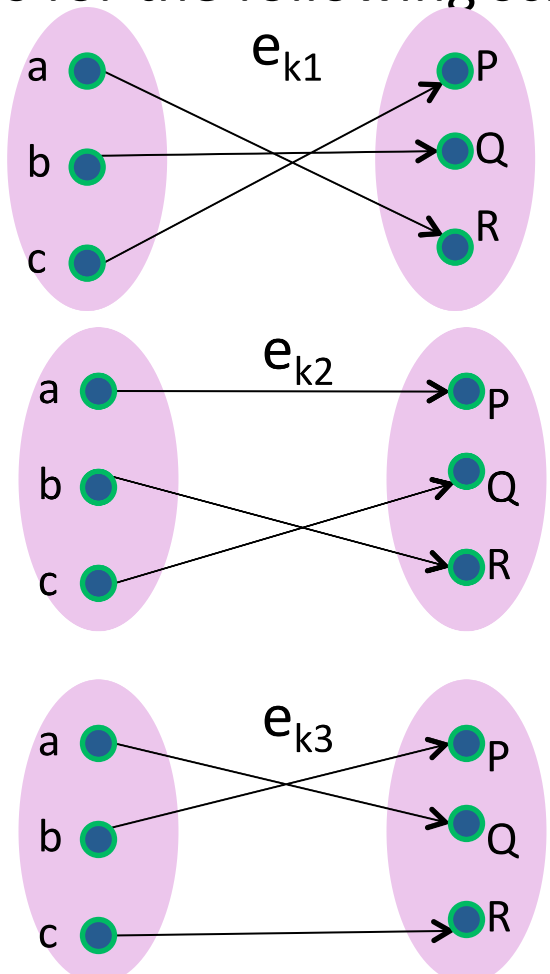
Intuitively, by seeing the safe, you learn nothing about what is in it

Perfect Secrecy Example

- Find the a posteriori probabilities for the following scheme
- Verify that it is perfectly secret.

plaintext
Pr[**X**=a] = 1/2
Pr[**X**=b] = 1/3
Pr[**X**=c] = 1/6

keyspace
Pr[**K**=k_1] = 1/3
Pr[**K**=k_2] = 1/3
Pr[K=k_3] = 1/3



Observations on Perfect Secrecy

Perfect Secrecy iff

$$\boxed{\Pr[Y = y \mid X = x] = \Pr[Y = y]}$$

Follows from
Baye's theorem

Perfect Indistinguishability

$\forall x_1, x_2 \in P$

$$\boxed{\Pr[Y = y \mid X = x_1] = \Pr[Y = y \mid X = x_2]}$$

Perfect secrecy has nothing to do with plaintext distribution.
Thus a crypto-scheme will achieve perfect secrecy irrespective of the language used in the plaintext.

Shift Cipher with a Twist

- Plaintext set : P = {0,1,2,3 …, 25}
- Ciphertext set : C = {0,1,2,3 …, 25}
- Keyspace : K = {0,1,2,3 …, 25}
- Encryption Rule : $e_K(x) = (x + K) \bmod 26$,
- Decryption Rule : $d_k(x) = (x - K) \bmod 26$

  where $K \in$ K and $x \in$ P

The Twist :

(1) the key changes after every encryption

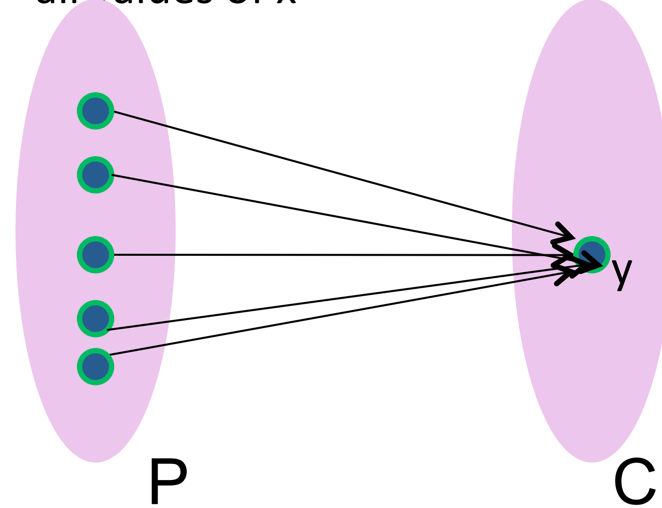
(2) keys are picked with uniform probability

The Twisted Shift Cipher is Perfectly Secure

$$\begin{aligned}
\mathbf{Pr}[\mathbf{y} = y] &= \sum_{K \in \mathbb{Z}_{26}} \mathbf{Pr}[\mathbf{K} = K]\mathbf{Pr}[\mathbf{x} = d_K(y)] \\
&= \sum_{K \in \mathbb{Z}_{26}} \frac{1}{26}\mathbf{Pr}[\mathbf{x} = y - K] \\
&= \frac{1}{26} \sum_{K \in \mathbb{Z}_{26}} \mathbf{Pr}[\mathbf{x} = y - K]. \\
&= \frac{1}{26}
\end{aligned}$$

Keys chosen with uniform probability

This is 1 because the sum is over all values of x



$$\begin{aligned}
\mathbf{Pr}[y|x] &= \mathbf{Pr}[\mathbf{K} = (y - x) \bmod 26] \\
&= \frac{1}{26}
\end{aligned}$$

For every pair of y and x, there is exactly one key . Probability of that key is 1/26

The Twisted Shift Cipher is Perfectly Secure

$$\begin{aligned}\mathbf{Pr}[\mathbf{y} = y] &= \sum_{K \in \mathbb{Z}_{26}} \mathbf{Pr}[\mathbf{K} = K]\mathbf{Pr}[\mathbf{x} = d_K(y)] \\ &= \sum_{K \in \mathbb{Z}_{26}} \frac{1}{26}\mathbf{Pr}[\mathbf{x} = y - K] \\ &= \frac{1}{26} \sum_{K \in \mathbb{Z}_{26}} \mathbf{Pr}[\mathbf{x} = y - K]. \\ &= \frac{1}{26}\end{aligned}$$

$$\boxed{\begin{aligned}\mathbf{Pr}[x|y] &= \frac{\mathbf{Pr}[x]\mathbf{Pr}[y|x]}{\mathbf{Pr}[y]} \\ &= \frac{\mathbf{Pr}[x]\frac{1}{26}}{\frac{1}{26}} \\ &= \mathbf{Pr}[x],\end{aligned}}$$

$$\begin{aligned}\mathbf{Pr}[y|x] &= \mathbf{Pr}[\mathbf{K} = (y - x) \bmod 26] \\ &= \frac{1}{26}\end{aligned}$$

Shannon's Theorem

If $|K| = |C| = |P|$ then the system provides perfect secrecy iff
(1) every key is used with equal probability $1/|K|$, and
(2) for every $x \in P$ and $y \in C$, there exists a unique key $k \in K$ such that $e_k(x) = y$

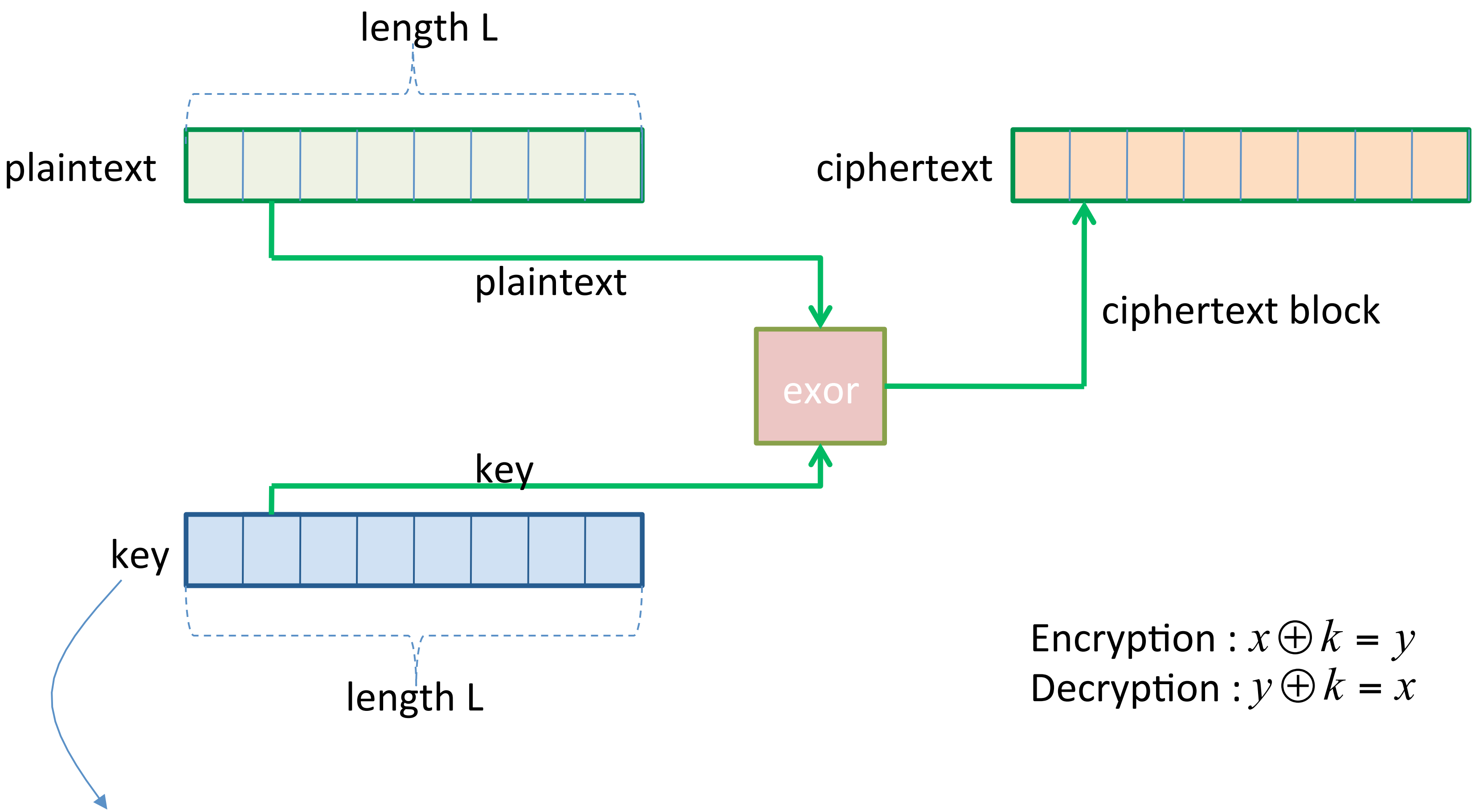
Intuition :

Every $y \in C$ can result from any of the possible plaintexts x

Since $|K| = |P|$ there is exactly one mapping from each plaintext to y

Since each key is equi-probable, each of these mappings is equally probable

One Time Pad (Verman's Cipher)



Encryption : $x \oplus k = y$
Decryption : $y \oplus k = x$

chosen uniformly from keyspace of size 2^L
$\Pr[\boldsymbol{K} = k] = 1/2^L$

One Tme Pad (Example)

e=000 h=001 i=010 k=011 l=100 r=101 s=110 t=111

Encryption: Plaintext ⊕ Key = Ciphertext

	h	e	i	l	h	i	t	l	e	r
Plaintext:	001	000	010	100	001	010	111	100	000	101
Key:	111	101	110	101	111	100	000	101	110	000
Ciphertext:	110	101	100	001	110	110	111	001	110	101
	s	r	l	h	s	s	t	h	s	r

CR

One Time Pad is Perfectly Secure

- Proof using indistinguishability

$$\Pr[Y=y \mid X=x] = \Pr[X=x, K=k \mid X=x] \qquad \text{from } x \oplus k = y$$

$$= \Pr[K=k] = \frac{1}{2^L}$$

$$\Pr[Y=y \mid X=x_1] = \frac{1}{2^L} = \Pr[Y=y \mid X=x_2]$$

$$\forall x_1, x_2 \in X$$

This implies perfect Indistinguishability that is independent of the plaintext distribution

Limitations of Perfect Secrecy

- Key must be at least as long as the message
  - Limits applicability if messages are long

- Key must be changed for every encryption
  - If the same key is used twice, then an adversary can compute the ex-or of the messages

$$x_1 \oplus k = y_1$$
$$x_2 \oplus k = y_2$$
$$x_1 \oplus x_2 = y_1 \oplus y_2$$

The attacker can then do language analysis to determine y_1 and y_2

Ciphers in Practice

- Perfect secrecy is difficult to achieve in practice
- Instead we use a crypto-scheme that cannot be ***broken in reasonable time*** with ***reasonable success***
- This means,
  - Security is only achieved against adversaries that run in polynomial time
  - Attackers can potentially succeed with a very small probability (attackers need to be very lucky to succeed)

Quantifying Information

A Metric to Quantify Information

There is one alphabet missing in each of these words. Can you find the alphabet so that the words make sense?

nough	enough
ntwork	network
hardwar	hardware

reate	create
shool	school
sott	scott

Frequently occurring letters (like e) contain less information than non-frequent letters (like c)

We need to have function to quantify information!

Additionally, the function should be (1) continuous (2) should be able to sum individual information (eg. X1 : Message 1, X2 : Message 2)

I(X1, X2) = I(Message 1) + I(Message 2)

Metric to Quantify Information

Claude Shannon

$$\log_2\left(\frac{1}{p_i}\right)$$

A higher probability indicates lesser information content.

Pr(e) = 0.12702	-log2(0.12702) = 2.97
Pr(a) = 0.08167	-log2(0.08167) = 3.61
Pr(m) = 0.02406	-log2(0.02406) = 5.37
Pr(c) = 0.02782	-log2(0.02782) = 5.16
Pr(q) = 0.0095	-log2(0.0095) = 6.71
...	
...	
...	

Metric to Quantify Information

Claude Shannon

To find the average information content of a language find weighted sum as follows

$$\sum_{i=1}^{n} p_i \log_2\left(\frac{1}{p_i}\right)$$

Metric to Quantify Information

Claude Shannon

To find the average information content of a language find weighted sum as follows

Call this term the **Entropy**

$$H(X) = \sum_{i=1}^{n} p_i \log_2\left(\frac{1}{p_i}\right)$$

Entropy provides the average number of bits needed to represent letters in the language

Entropy of English

Contemporary : 4.03 bits

Shakesphere : 4.106 bits

German : 4.08 bits

French : 4.00 bits

Italian : 3.98 bits

Spanish : 3.98 bits

Maximum Entropy occurs when each alphabet is equally likely (ie. 1/26).
The maximum entropy is -log_2(1/26) = 4.7

CR

Entropy of the Weather Forecast

Weather Forecast

Tomorrow I the weather will be ___________

M1 : Sunny (with probability 0.05) M2 : Cloudy (with probability 0.15) M3 : Light Rain (with probability 0.70) M4 : Heavy Rain (with probability 0.10)

$$H(Forecast) = \sum_{i=1}^{n} p_i \log_2\left(\frac{1}{p_i}\right)$$

$$= -((0.05)\log_2 0.05 + (0.15)\log_2 0.15 + (0.7)\log_2 0.7 + (0.1)\log_2 0.1)$$

$$= 1.319$$

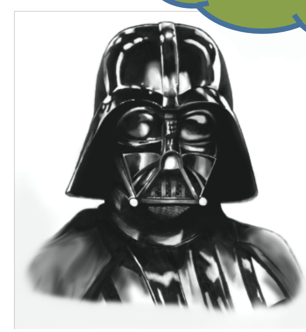
CR

Entropy and Uncertainity

- Alice thinks of a number (0 or 1)
- The choice is denoted by a discrete random variable X.

X

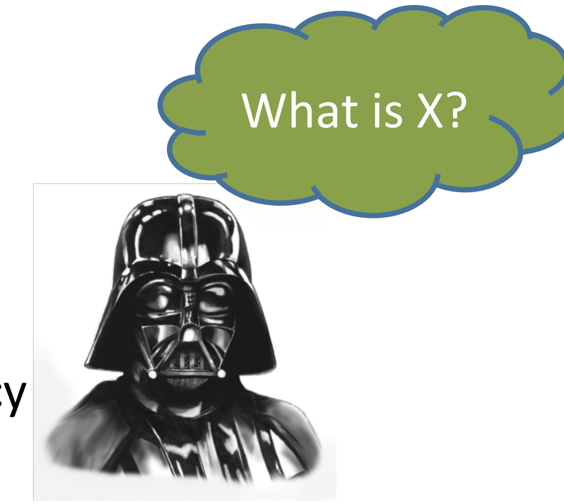
What is X?

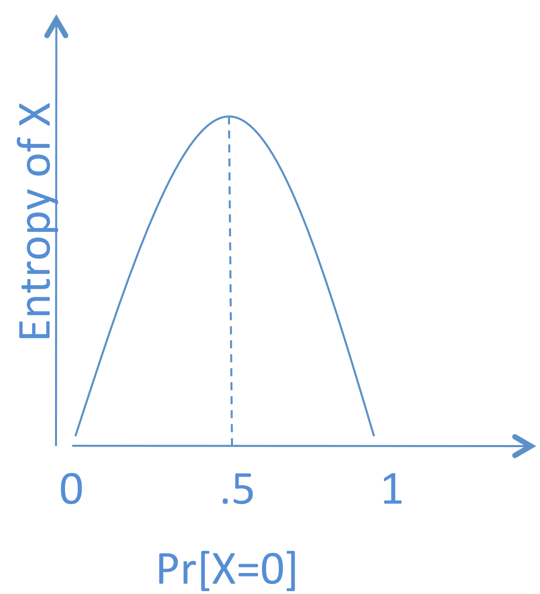


- What is Mallory’s uncertainty about X?
  - Depends on the probability distribution of X

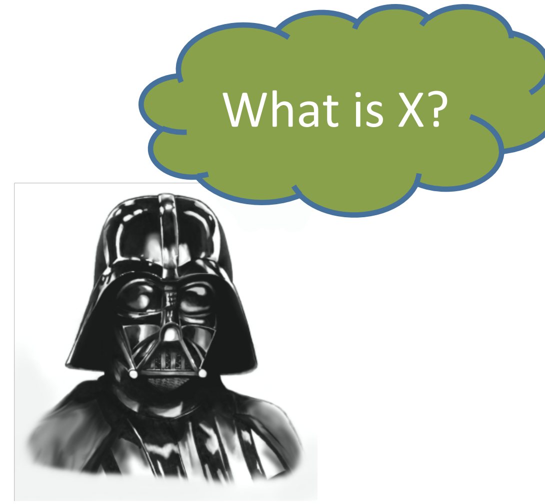
Uncertainty

- Lets assume Mallory know this probability distribution.
- If Pr[X = 1] = 1 and Pr[X = 0] = 0
  - Then Mallory can determine with 100% accuracy
- If Pr[X = 0] = .75 and Pr[X = 1] = .25
  - Mallory will guess X as 0, and gets it right 75% of the time
- If Pr[X=0] = Pr[X = 1] = 0.5
  - Mallory's guess would be similar to a uniformly random guess. Gets it right ½ the time.





What is the Entropy of X?

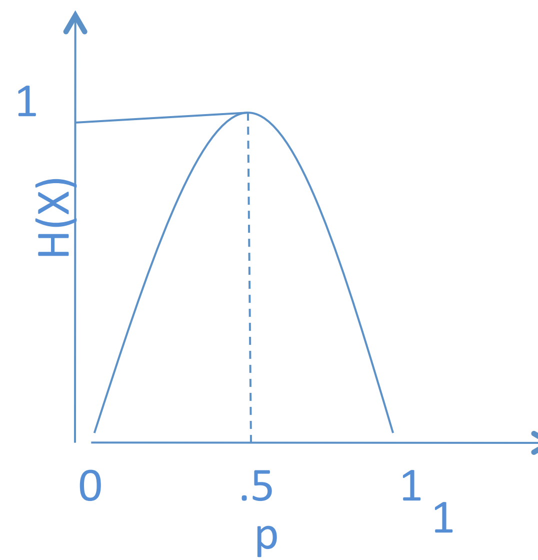


Pr[X=0] = p and Pr[X=1] = 1 - p

$H(X) = -\, p\log_2 p - (1-p)\,\log_2(1-p)$

$H(X)_{p=0} = 0,\ H(X)_{p=1} = 0,\ H(X)_{p=.5} = 1$

using $\lim_{p\to 0}(p \log\ p) = 0$



Properties of H(X)

- If X is a random variable, which takes on values {1,2,3,….n} with probabilities p_1, p_2, p_3, ….p_n, then

1. $H(X) \leq \log_2 n$

2. When $p_1 = p_2 = p_3 = \ldots p_n = 1/n$ then $H(X) = \log_2 n$

   Example an 8 face dice.
   If the dice is fair, then we obtain the maximum entropy of 3 bits
   If the dice is unfair, then the entropy is < 3 bits

Entropy and Coding

- Entropy quantifies Information content

  "Can we encode a message M in such a way that the average length is as short as possible and hopefully equal to H(M)?"

  Huffman Codes :

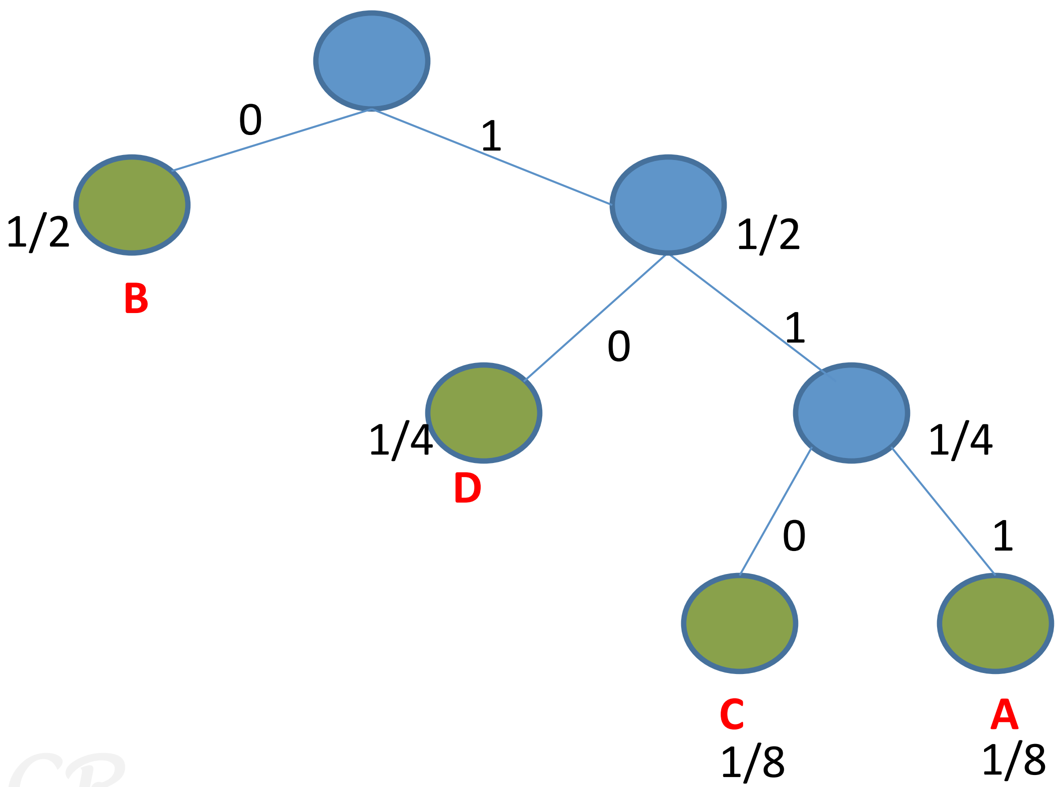
  allocate more bits to least probable events
  allocate less bits to popular events

Example

- S = {A, B, C, D} are 4 symbols
- Probability of Occurrence is :

  P(A) = 1/8, P(B) = ½, P(C) = 1/8, P(D) = 1/4

Encoding
A : 111
B : 0
C : 110
D: 10



To decode, with each bit traverse the tree from root until you reach a leaf.

Decode this?
1101010111

Example :
Average Length and Entropy

- S = {A, B, C, D} are 4 symbols
- Probability of Occurrence is :

  p(A) = 1/8, p(B) = ½, p(C) = 1/8, p(D) = ¼

Encoding
A : 111
B : 0
C : 110
D: 10

- Average Length of Huffman code :

  3*p(A) + 1*p(B) + 3*p(C) + 2*p(D) = 1.75

- Entropy H(S) =

  $-1/8 \log_2(8) - ½ \log_2(2) - 1/8 \log_2(8) - ¼ \log_2(4)$

  = 1.75

Measuring the Redundancy in a Language

- Let S be letter in a language (eg. S = {A,B,C,D})
- $\mathbf{S} = S \times S \times S \times S \times S \times S\,(k\ \text{times})$ is a set representing messages of length k
- Let $S^{(k)}$ be a random variable in $\mathbf{S}$
- The average information in each letter is given by the rate of $S^{(k)}$.

$$r_k = \frac{H(S^{(k)})}{k}$$

- r_k for English is between 1.0 and 1.5 bits/letter

Measuring the Redundancy in a Language

- Absolute Rate : The maximum amount of information per character in a language
  - the absolute rate of language S is $R = \log_2 |S|$
  - For English, $|S| = 26$, therefore $R = 4.7$ bits / letter

- Redundancy of a language is

$$D = R - r_k$$

  - For English when rk = 1, then D = 3.7 → around 70% redundant

Example (One letter analysis)

- Consider a language with 26 letters of the set S = {s_1, s_2, s_3, ….., s_{26}}. Suppose the language is characterized by the following probabilities. **What is the language redundancy?**

$$P(s_1) = \frac{1}{2},\ P(s_2) = \frac{1}{4}$$

$$P(s_i) = \frac{1}{64} \quad for \quad i = 3,4,5,6,7,8,9,10$$

$$P(s_i) = \frac{1}{128} \quad for \quad i = 11,12,\dots,26$$

Absolute Rate

$$R = \log 26 = 4.7$$

Rate of the Language for 1 letter analysis

$$\begin{aligned} r_1 &= H(S^{(1)}) \\ &= \sum_{i=1}^{26} P(s_i) \log \frac{1}{P(s_i)} \\ &= \frac{1}{2}\log 2 + \frac{1}{4}\log 4 + 8\left(\frac{1}{64}\log 64\right) + 16\left(\frac{1}{128}\log 128\right) \\ &= \frac{1}{2} + \frac{1}{2} + \frac{6}{8} + \frac{7}{8} = 2.625 \end{aligned}$$

Language Redundancy

$$D = R - r_1 = 4.7 - 2.625 = 2.075$$

Language is ~70% redundant

Example (Two letter analysis)

- In the set S = {s_1, s_2, s_3, ….., s_{26}}, suppose the diagram probabilites is as below. **What is the language redundancy?**

$$P(s_{i+1} \mid s_i) = P(s_{i+2} \mid si) = \frac{1}{2} \quad for \quad i = 1 \ to \ 24$$

$$P(s_{26} \mid s_{25}) = P(s_1 \mid s_{25}) = P(s_1 \mid s_{26}) = P(s_2 \mid s_{26}) = \frac{1}{2}$$

$$all \ \ other \ \ probabilities \ \ are \ \ 0$$

$$P(s_1, s_2) = P(s_2 \mid s_1) \times P(s_1) = 1/4; \ P(s_1, s_3) = P(s_3 \mid s_1) \times P(s_1) = 1/4$$

$$P(s_2, s_3) = P(s_3 \mid s_2) \times P(s_2) = 1/8; \ P(s_2, s_4) = P(s_4 \mid s_2) \times P(s_2) = 1/8$$

$$P(s_i, s_{i+1}) = P(s_{i+1} \mid s_i) P(s_i) = 1/128 \quad for \ \ i = 3, 4, \ldots\ldots, 10$$

$$P(s_i, s_{i+2}) = P(s_{i+2} \mid s_i) P(s_i) = 1/128 \quad for \ \ i = 3, 4, \ldots\ldots, 10$$

$$P(s_i, s_{i+1}) = P(s_{i+1} \mid s_i) P(s_i) = 1/256 \quad for \ \ i = 11, 12, \ldots\ldots, 24$$

$$P(s_i, s_{i+2}) = P(s_{i+2} \mid s_i) P(s_i) = 1/256 \quad for \ \ i = 11, 12, \ldots\ldots, 24$$

$$P(s_{25}, s_{26}) = P(s_{25}, s_1) = P(s_{26}, s_1) = P(s_{26}, s_2) = 1/256$$

Rate of the Language for 2 letter analysis

$$r_2 = H(S^{(2)})/2$$

$$= \frac{1}{2} \sum_{i,j=1}^{26} P(s_i, s_j) \log \frac{1}{P(s_i, s_j)}$$

$$= \frac{1}{2} \left[2\left(\frac{1}{4} \log 4\right) + 2\left(\frac{1}{8} \log 8\right) + 16\left(\frac{1}{128} \log 128\right) + 32\left(\frac{1}{256} \log 256\right) \right]$$

$$= \frac{1}{2} \left[1 + \frac{3}{4} + \frac{7}{8} + 1 \right] = \frac{3.625}{2} = 1.8125$$

Language Redundancy

$$D = R - r_2 = 4.7 - 1.8125 = 2.9$$

Language is ~60% redundant

CR

Observations

$$\text{Single letter analysis} : r_1 = H(S^{(1)}) = 2.625; \;\; D = 2.075$$

$$\text{Two letter analysis} : H(S^{(2)}) = 3.625; \;\; r_2 = 1.8125; \;\; D = 2.9$$

- H(S(2)) – H(S(1)) = 1 bit
  - why?
- As we increase the message size
  - Rate reduces; inferring less information per letter
  - Redundancy increases

Conditional Entropy

- Suppose X and Y are two discrete random variables, then conditional entropy is defined as

$$H(X \mid Y) = \sum_{y} p(y) \sum_{x} p(x \mid y) \log_2 \left(\frac{1}{p(x \mid y)} \right)$$

$$= \sum_{x} \sum_{y} p(x, y) \log_2 \left(\frac{p(x)}{p(x, y)} \right)$$

- Conditional entropy means ….
  - What is the remaining uncertainty about X given Y
  - $H(X|Y) \leq H(X)$ with equality when X and Y are independent

Derive using the fact that $p(a|b) = p(a,b) / p(b)$

Joint Entropy

- Suppose X and Y are two discrete random variables, and p(x,y) the value of the joint probability distribution when X=x and Y=y
- Then the joint entropy is given by

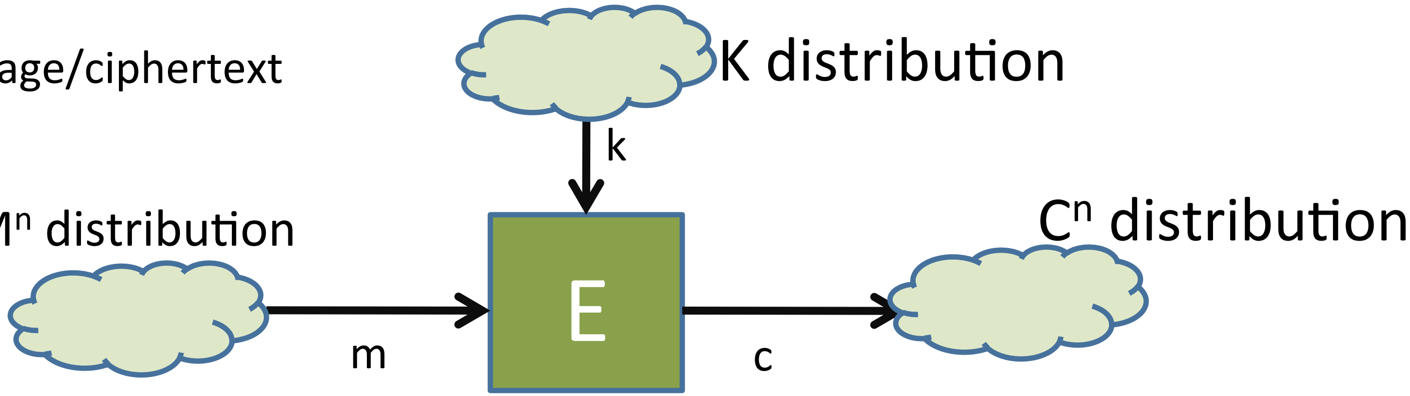
$$H(X,Y) = \sum_{y} \sum_{x} p(x,y) \log_2 \left(\frac{1}{p(x,y)} \right)$$

- The joint entropy is the average uncertainty of 2 random variables

CR

Entropy and Encryption

n: length of message/ciphertext



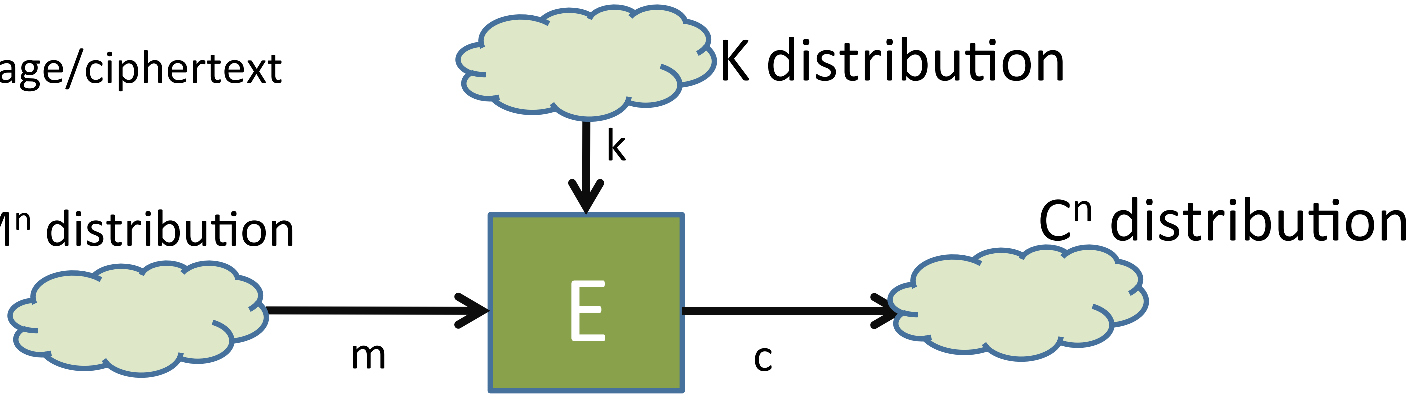
- There are three entropies: $H(P^{(n)})$, $H(K)$, $H(C^{(n)})$
- Message Equivocation :

  If the attacker can view n ciphertexts, what is his uncertainty about the message

$$H(M^{(n)} \mid C^{(n)}) = \sum_{c \in C^n} p(c) \sum_{m \in M^n} p(m \mid c) \log_2\left(\frac{1}{p(m \mid c)}\right)$$

Entropy and Encryption

n: length of message/ciphertext



- Key Equivocation :

  If the attacker can view n ciphertexts, what is his uncertainty about the key

$$H(K \mid C^{(n)}) = \sum_{c \in C^n} p(c) \sum_{m \in M^n} p(k \mid c) \log_2\left(\frac{1}{p(k \mid c)}\right)$$

Unicity Distance

$$H(K \mid C^{(n)}) = \sum_{c \in C^n} p(c) \sum_{m \in M^n} p(k \mid c) \log_2\left(\frac{1}{p(k \mid c)}\right)$$

- As n increases, $H(K|C^{(n)})$ reduces…
  - This means that the uncertainty of the key reduces as the attacker observes more ciphertexts
- Unicity distance is the value of n for which $H(K \mid C^{(n)}) \approx 0$
  - This means, the entire key can be determined in this case

CR

Unicity Distance and Classical Ciphers

Cipher	Unicity Distance (for English)
Caesar's Cipher	1.5 letters
Affine Cipher	2.6 letters
Simple Substitution Cipher	27.6 letters
Permutation Cipher	0.12 (block size = 3) 0.66 (block size = 4) 1.32 (block size = 5) 2.05 (block size = 6)
Vigenere Cipher	1.47d (d is the key length)

Product Ciphers

- Consider a cryptosystem where P=C (this is an endomorphic system)
  - Thus the ciphertext and the plaintext set is the same
- Combine two ciphering schemes to build a product cipher

Given two endomorphic crypto-systems

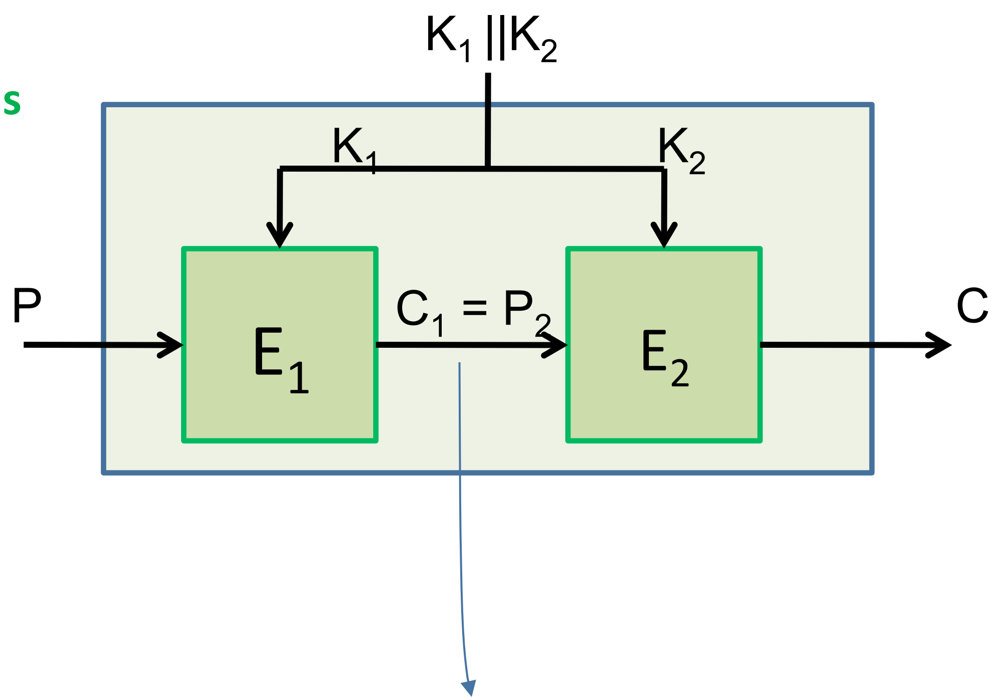
$$S_1 : (P, P, K_1, E_1, D_1)$$

$$S_2 : (P, P, K_2, E_2, D_2)$$

Resultant Product Cipher

$$S_1 \times S_2 : (P, P, K_1 \times K_2, E, D)$$

Resultant Key Space $K_1 \times K_2$



Ciphertext of first cipher fed as input to the second cipher

Product Ciphers

- Consider a cryptosystem where P=C (this is an endomorphic system)
  - Thus the ciphertext and the plaintext set is the same
- Combine two ciphering schemes to build a product cipher

Given two endomorphic crypto-systems

$$S_1 : x = d_{K_1}(e_{K_1}(x))$$

$$S_2 : x = d_{K_2}(e_{K_2}(x))$$

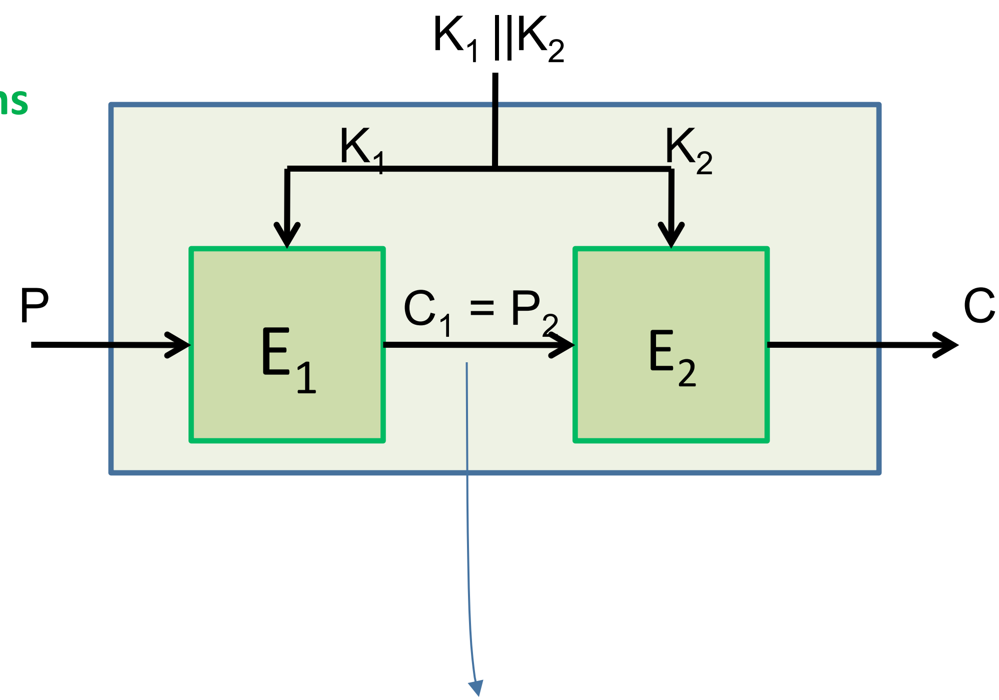
Resultant Product Cipher

$$S_1 \times S_2$$

$$e_{(K_1,K_2)}(x) = e_{K_2}(e_{K_1}(x))$$

$$d_{(K_1,K_2)}(x) = d_{K_2}(d_{K_1}(x))$$

Resultant Key Space $K_1 \times K_2$



Ciphertext of first cipher fed as input to the second cipher

Affine Cipher is a Product Cipher

- P = C = {0, 1, 2, … 25}

Multiplicative Cipher

Shift Cipher

Affine Cipher = M x S

Encryption ($e_a(x)$) : $y = ax \bmod 26$
Decryption ($d_a(x)$) : $x = a^{-1}y \bmod 26$

Encryption ($e_b(x)$) : $y = x+b \bmod 26$
Decryption ($d_b(x)$) : $x = y-b \bmod 26$

- Affine cipher : $y = ax + b \bmod 26$
- Size of Key space is
  - Size of key space for Multiplicative cipher * Size of keyspace for shift cipher
  - 12 * 26 = 312

Is S x M same as the Affine Cipher

- S x M : $y = a(x + b) \bmod 26$
  $= ax + ba \bmod 26$
- Key is (b,a)
- ba mod 26 is some b' such that
  $a^{-1}b' = b \bmod 26$
- This can be represented as an Affine cipher,
  $$y = ax + b' \bmod 26$$

Thus affine ciphers are commutable (i.e. S x M = M x S)

Create a non-commutable product ciphers

CR

Idempotent Ciphers

- If $S_1 : (P, P, K, E_1, D_1)$ is an endomorphic cipher
- then it is possible to construct product ciphers of the form S_1 x S_1, denoted $S^2 \; : (P, P, K \times K, E, D)$
- If $S^2 = S$ then the cipher is called idempotent cipher

Show that the simple substitution cipher is idempotent

Does the security of the newly formed cipher increase?

In a non-idempotent cipher, however the security may increase.

Iterative Cipher

- An n-fold product of this is **S x S x S … (n times) = S^n** is an iterative cipher

  All modern block ciphers like DES, 3-DES, AES, etc. are iterative, non-idempotent, product ciphers.

  We will see more about these ciphers next!!