Information Theory

Mohamed Hamada

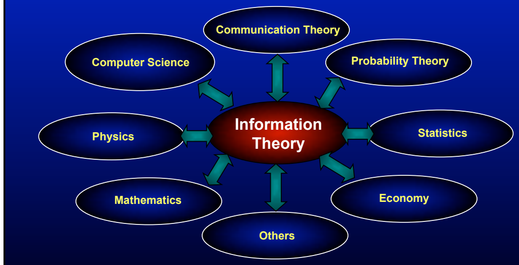
Software Engineering Lab
The University of Aizu

Email: hamada@u-aizu.ac.jp
URL: http://www.u-aizu.ac.jp/~hamada

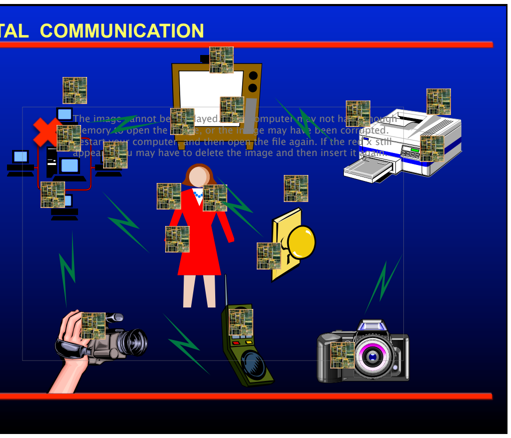
Today’s Topics

- Overview of Information Theory
- Digital Communication
- History
- What is Information Theory

OVERVIEW OF INFORMATION THEORY FRAMEWORK



DIGITAL COMMUNICATION



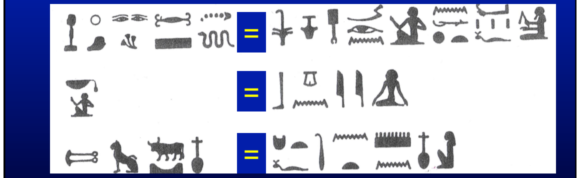
DIGITAL COMMUNICATION

DIGITAL COMMUNICATION

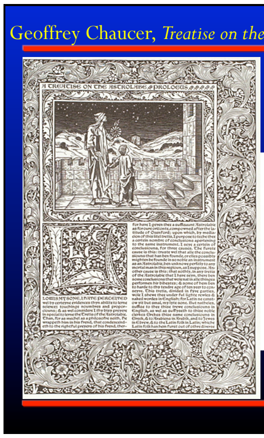
History

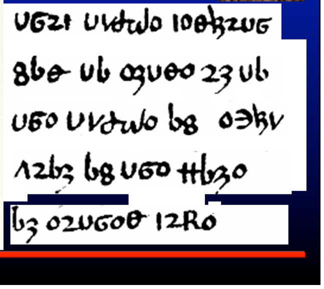
EGYPTIAN

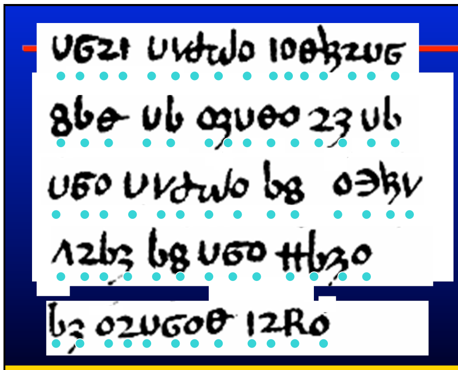
Cryptography, ca. 1900BC



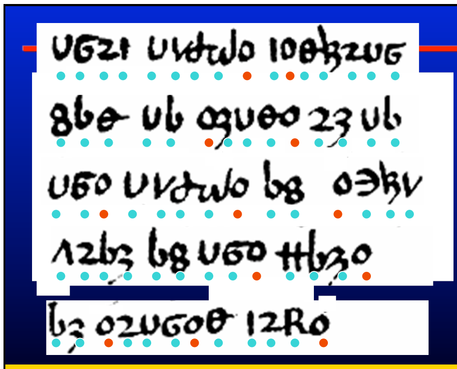
Geoffrey Chaucer, *Treatise on the Astrolabe,* 1391



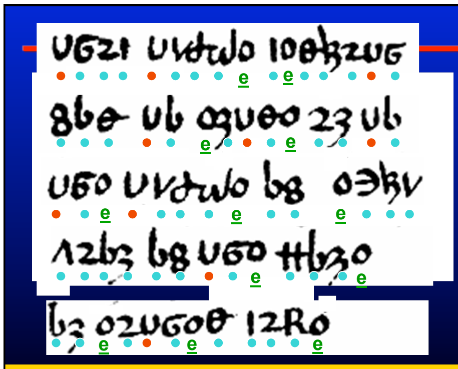




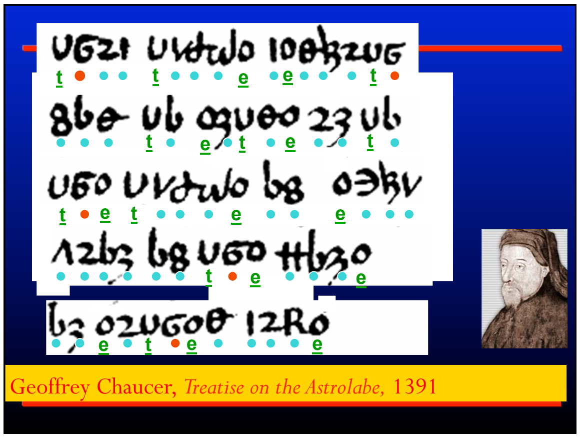
Geoffrey Chaucer, *Treatise on the Astrolabe,* 1391



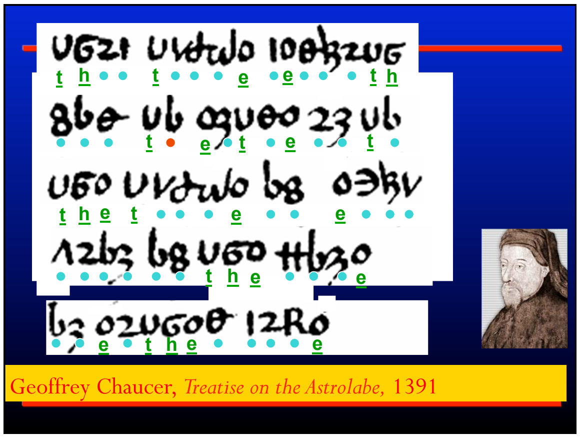
Geoffrey Chaucer, *Treatise on the Astrolabe,* 1391


e
e
e
e
e
e
e
e
e
e
e
e
e

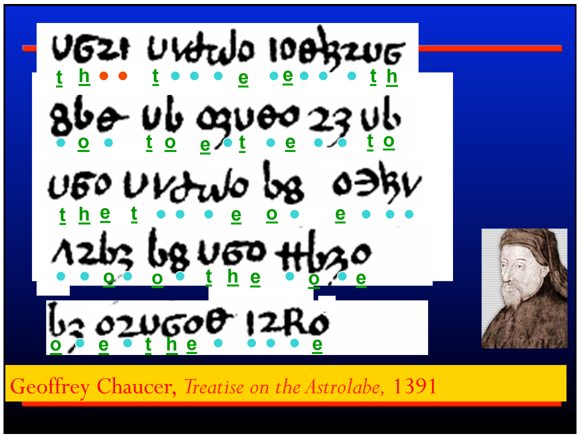
Geoffrey Chaucer, *Treatise on the Astrolabe,* 1391



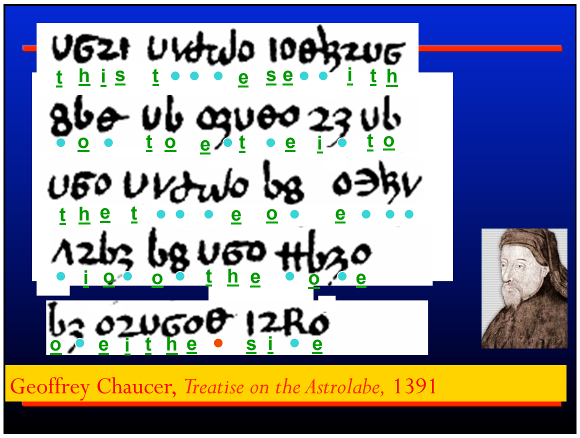
Geoffrey Chaucer, *Treatise on the Astrolabe,* 1391



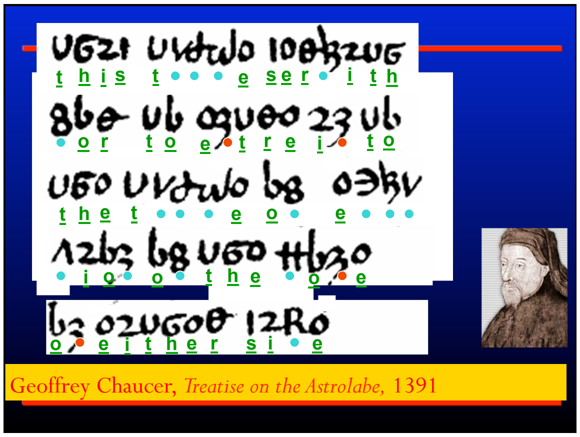
Geoffrey Chaucer, *Treatise on the Astrolabe,* 1391



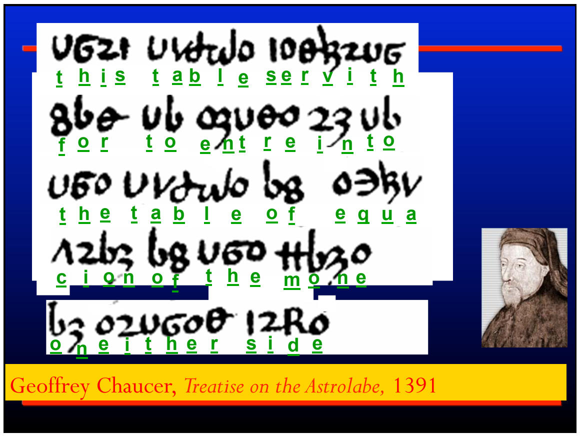
Geoffrey Chaucer, *Treatise on the Astrolabe,* 1391



Geoffrey Chaucer, *Treatise on the Astrolabe,* 1391

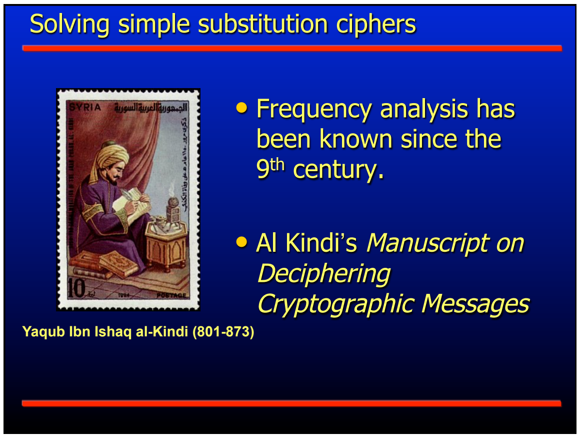


Geoffrey Chaucer, *Treatise on the Astrolabe,* 1391



this table serveth
for to entre into
the table of equa
cion of the mone
on either side

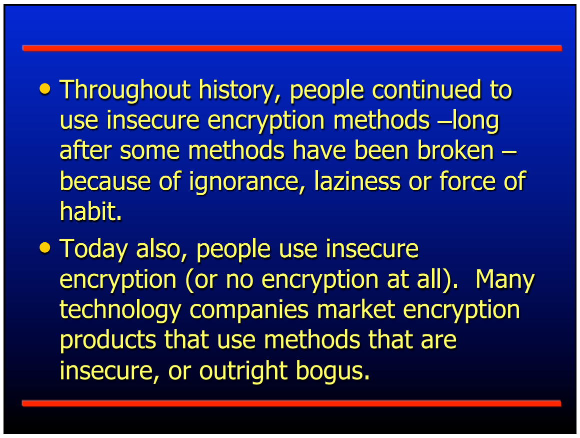
Geoffrey Chaucer, *Treatise on the Astrolabe,* 1391



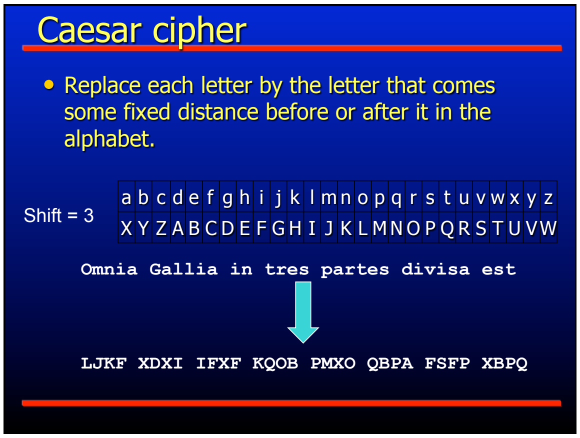
Solving simple substitution ciphers

Yaqub Ibn Ishaq al-Kindi (801-873)

- Frequency analysis has been known since the 9th century.
- Al Kindi's *Manuscript on Deciphering Cryptographic Messages*



- Throughout history, people continued to use insecure encryption methods –long after some methods have been broken – because of ignorance, laziness or force of habit.
- Today also, people use insecure encryption (or no encryption at all). Many technology companies market encryption products that use methods that are insecure, or outright bogus.



Caesar cipher

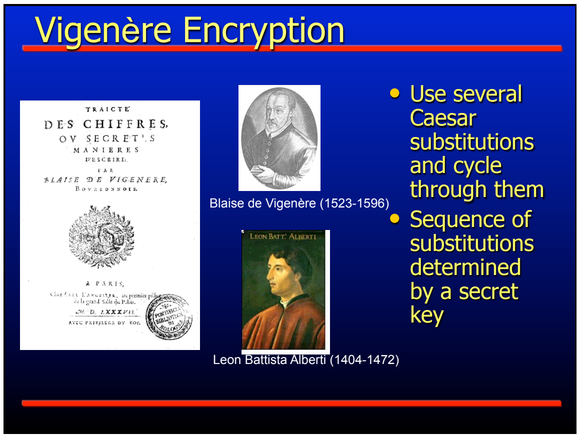
- Replace each letter by the letter that comes some fixed distance before or after it in the alphabet.

Shift = 3

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

```
Omnia Gallia in tres partes divisa est
```

```
LJKF XDXI IFXF KQOB PMXO QBPA FSFP XBPQ
```



Vigenère Encryption

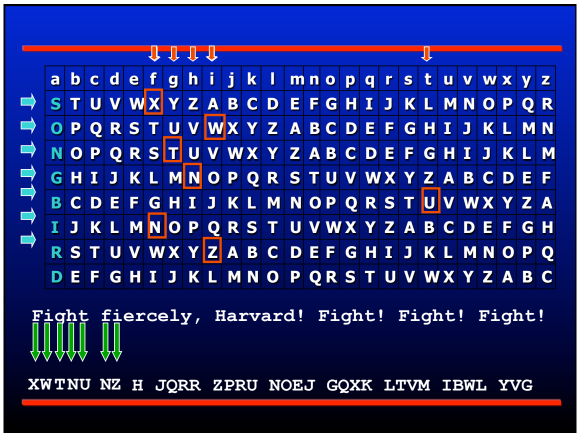
TRAICTÉ
DES CHIFFRES,
OV SECRET'.S
MANIERES
D'ESCRIRE.
PAR
BLAISE DE VIGENERE,
BOVRBONNOIS.

A PARIS,
Chez Abel L'Angelier, au premier pillier
de la grand' Salle du Palais.
M. D. LXXXVII.
AVEC PRIVILEGE DV ROY.

Blaise de Vigenère (1523-1596)

Leon Battista Alberti (1404-1472)

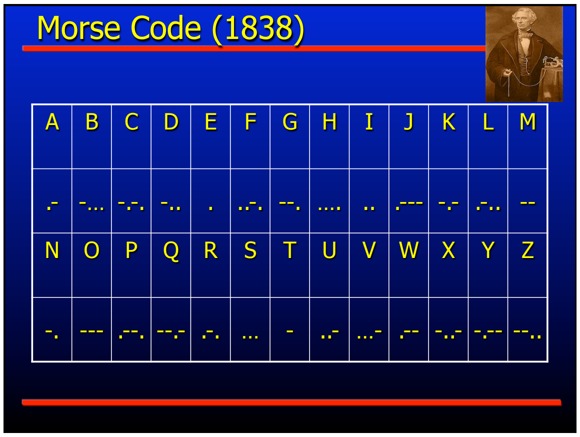
- Use several Caesar substitutions and cycle through them
- Sequence of substitutions determined by a secret key



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

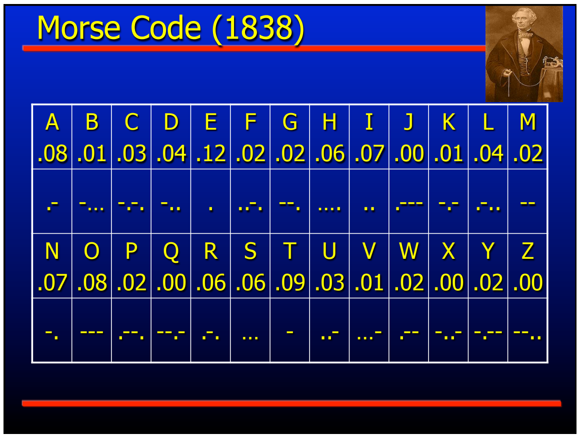
```
Fight fiercely, Harvard! Fight! Fight! Fight!
```

```
XWTNU NZ H JQRR ZPRU NOEJ GQXK LTVM IBWL YVG
```



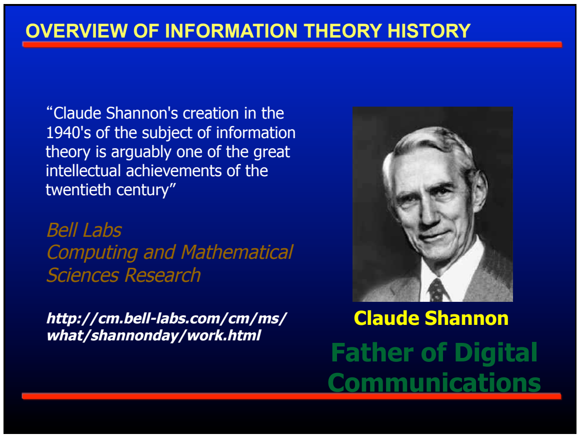
Morse Code (1838)

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
-.	---	.--.	--.-	.-.	...	-	..-	...-	.--	-..-	-.--	--..



Morse Code (1838)

A .08	B .01	C .03	D .04	E .12	F .02	G .02	H .06	I .07	J .00	K .01	L .04	M .02
.-	-...	-.-.	-..	.	..-.	--.	....	..	.---	-.-	.-..	--
N .07	O .08	P .02	Q .00	R .06	S .06	T .09	U .03	V .01	W .02	X .00	Y .02	Z .00
-.	---	.--.	--.-	.-.	...	-	..-	...-	.--	-..-	-.--	--..



OVERVIEW OF INFORMATION THEORY HISTORY

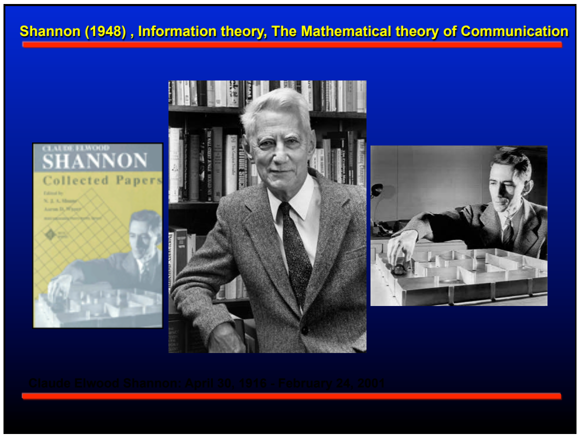
"Claude Shannon's creation in the 1940's of the subject of information theory is arguably one of the great intellectual achievements of the twentieth century"

Bell Labs
Computing and Mathematical Sciences Research

http://cm.bell-labs.com/cm/ms/what/shannonday/work.html

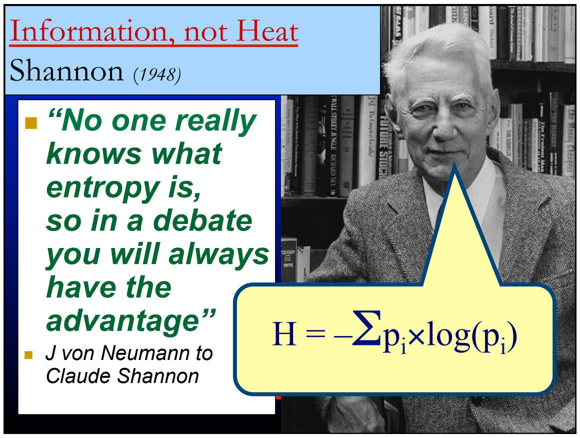
Claude Shannon

Father of Digital Communications



Shannon (1948) , Information theory, The Mathematical theory of Communication

Claude Elwood Shannon: April 30, 1916 - February 24, 2001

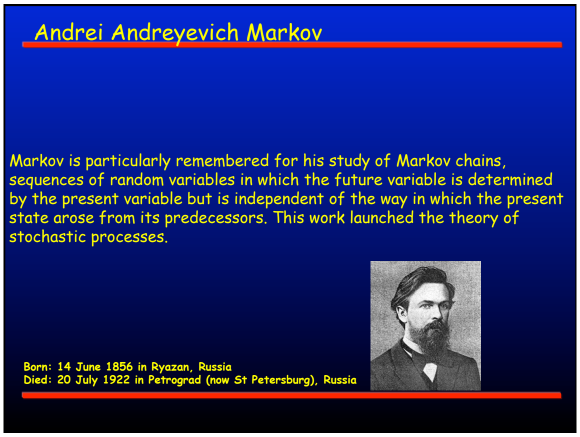


Information, not Heat

Shannon *(1948)*

- ***"No one really knows what entropy is, so in a debate you will always have the advantage"***
- *J von Neumann to Claude Shannon*

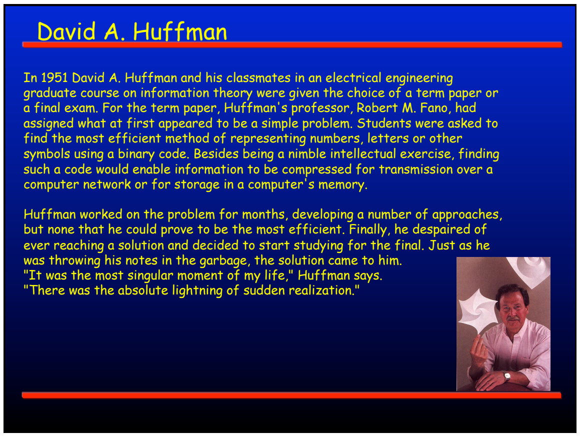
$$H = -\Sigma p_i \times \log(p_i)$$



Andrei Andreyevich Markov

Markov is particularly remembered for his study of Markov chains, sequences of random variables in which the future variable is determined by the present variable but is independent of the way in which the present state arose from its predecessors. This work launched the theory of stochastic processes.

Born: 14 June 1856 in Ryazan, Russia
Died: 20 July 1922 in Petrograd (now St Petersburg), Russia

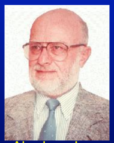


David A. Huffman

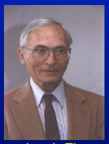
In 1951 David A. Huffman and his classmates in an electrical engineering graduate course on information theory were given the choice of a term paper or a final exam. For the term paper, Huffman's professor, Robert M. Fano, had assigned what at first appeared to be a simple problem. Students were asked to find the most efficient method of representing numbers, letters or other symbols using a binary code. Besides being a nimble intellectual exercise, finding such a code would enable information to be compressed for transmission over a computer network or for storage in a computer's memory.

Huffman worked on the problem for months, developing a number of approaches, but none that he could prove to be the most efficient. Finally, he despaired of ever reaching a solution and decided to start studying for the final. Just as he was throwing his notes in the garbage, the solution came to him.
"It was the most singular moment of my life," Huffman says.
"There was the absolute lightning of sudden realization."

The inventors



Abraham Lempel



Jacob Ziv

LZW (**Lempel-Ziv-Welch**) is an implementation of a lossless data compression algorithm created by Lempel and Ziv. It was published by Terry Welch in 1984 as an improved version of the LZ78 dictionary coding algorithm developed by Abraham Lempel and Jacob Ziv.

Fano

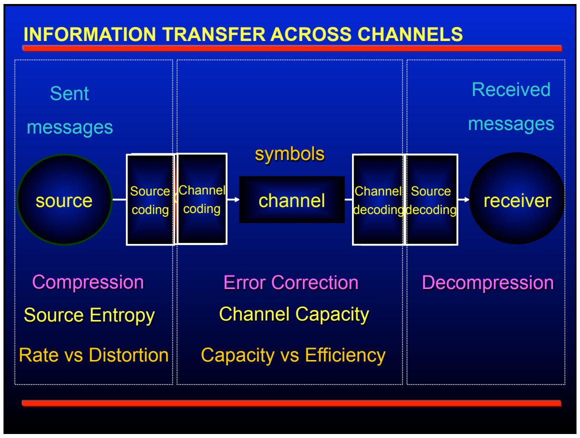
Shannon showed that it is possible to compress information. He produced examples of such codes which are now known as Shannon-Fano codes.

Robert Fano was an electrical engineer at MIT (the son of G. Fano, the Italian mathematician who pioneered the development of finite geometries and for whom the Fano Plane is named).

Low and High Information Content Messages

- The more frequent a message is, the less information it conveys when it occurs
- Two weather forecast messages: ☼ ☁
- Bos: ☼ ☁ ☁ ☼ ☼ ☁ ☁ ☁ ☼
- LA: ☼ ☼ ☼ ☼ ☼ ☼ ☼ ☁ ☼
- In LA "Sunny" is a low information message and "cloudy" is a high information message

INFORMATION TRANSFER ACROSS CHANNELS



Sent messages

source

Source coding

Channel coding

symbols

channel

Channel decoding

Source decoding

receiver

Received messages

Compression

Source Entropy

Rate vs Distortion

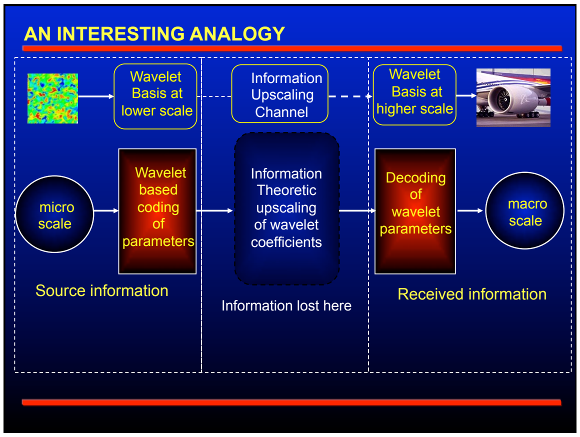
Error Correction

Channel Capacity

Capacity vs Efficiency

Decompression

AN INTERESTING ANALOGY



Wavelet Basis at lower scale

Information Upscaling Channel

Wavelet Basis at higher scale

micro scale

Wavelet based coding of parameters

Information Theoretic upscaling of wavelet coefficients

Decoding of wavelet parameters

macro scale

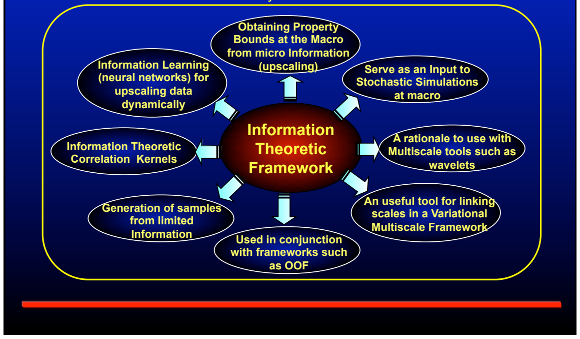
Source information

Information lost here

Received information

Applications of information theory with multiscale methods

Some currently ongoing and envisaged applications of Information Theory in a Multiscale Framework



Information Theoretic Framework

Obtaining Property Bounds at the Macro from micro Information (upscaling)

Serve as an Input to Stochastic Simulations at macro

A rationale to use with Multiscale tools such as wavelets

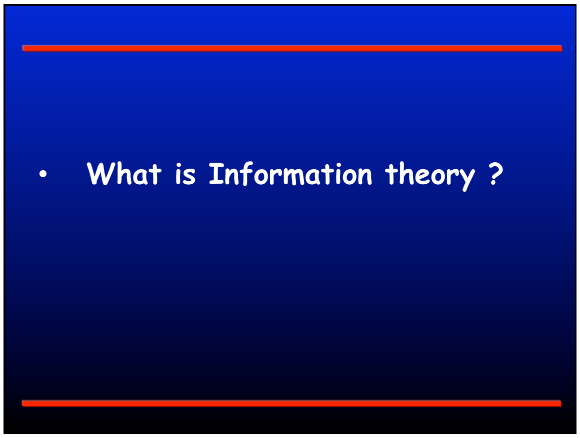
An useful tool for linking scales in a Variational Multiscale Framework

Used in conjunction with frameworks such as OOF

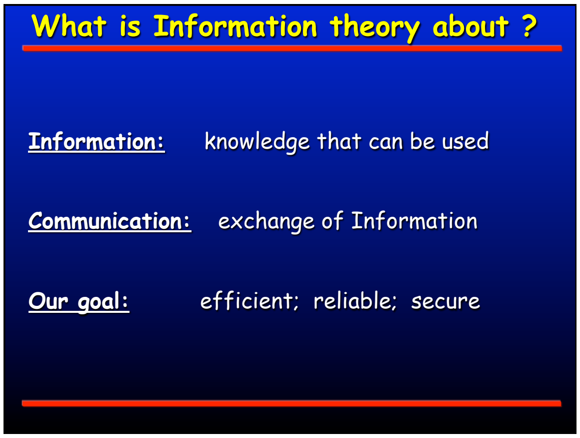
Generation of samples from limited Information

Information Theoretic Correlation Kernels

Information Learning (neural networks) for upscaling data dynamically



- What is Information theory ?

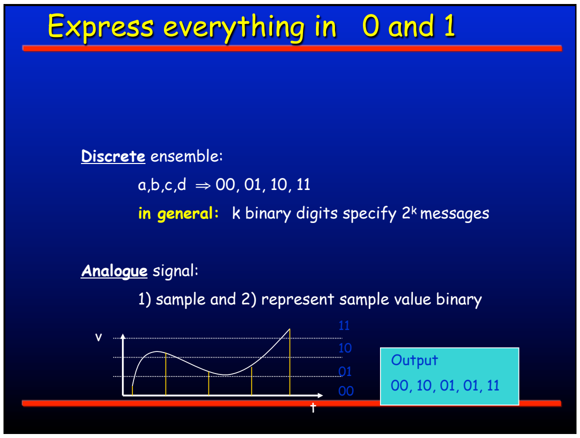


What is Information theory about ?

Information: knowledge that can be used

Communication: exchange of Information

Our goal: efficient; reliable; secure



Express everything in 0 and 1

Discrete ensemble:

a,b,c,d ⇒ 00, 01, 10, 11

in general: k binary digits specify 2^k messages

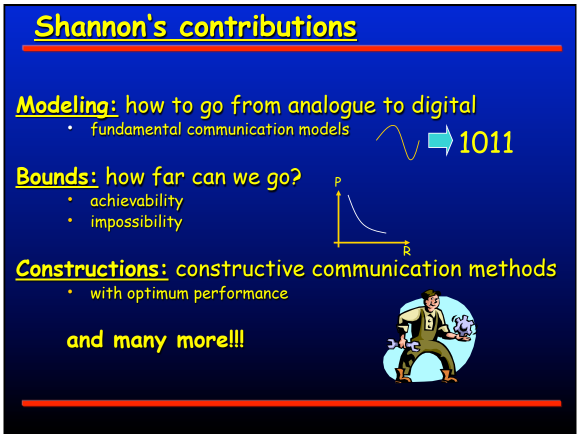
Analogue signal:

1) sample and 2) represent sample value binary

v
11
10
01
00
t

Output

00, 10, 01, 01, 11



Shannon's contributions

Modeling: how to go from analogue to digital

- fundamental communication models

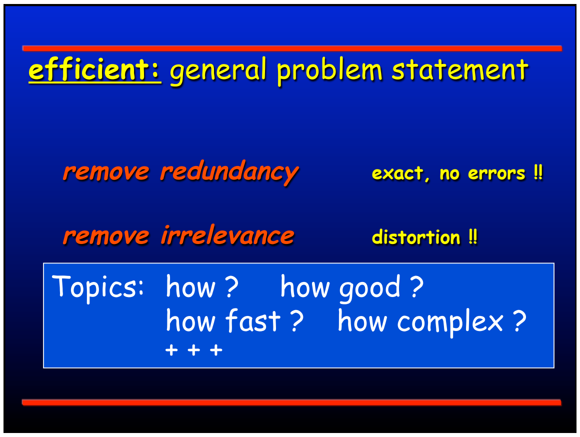
Bounds: how far can we go?

- achievability
- impossibility

Constructions: constructive communication methods

- with optimum performance

and many more!!!

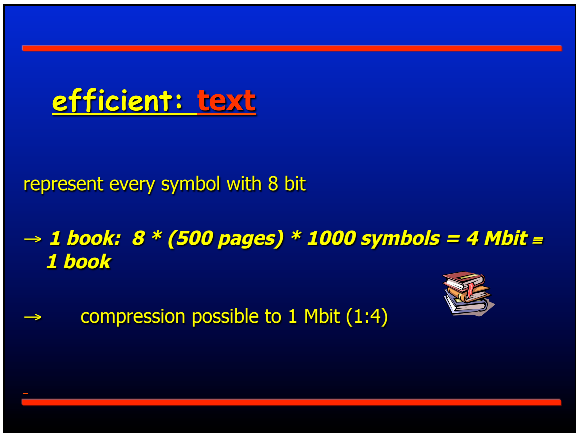


efficient: general problem statement

remove redundancy **exact, no errors !!**

remove irrelevance **distortion !!**

Topics: how ? how good ?
how fast ? how complex ?
+ + +

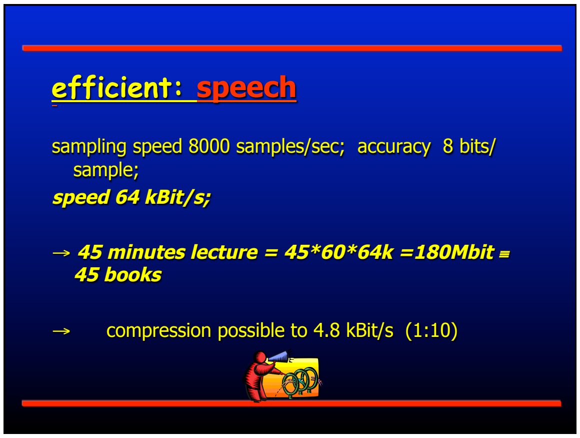


efficient: text

represent every symbol with 8 bit

→ ***1 book: 8 * (500 pages) * 1000 symbols = 4 Mbit = 1 book***

→ compression possible to 1 Mbit (1:4)



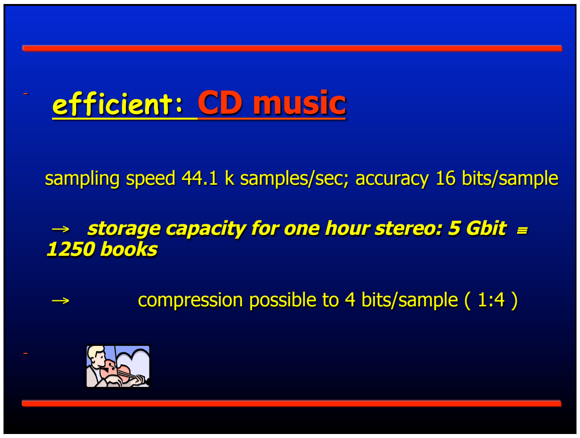
efficient: speech

sampling speed 8000 samples/sec; accuracy 8 bits/ sample;

speed 64 kBit/s;

→ ***45 minutes lecture = 45*60*64k =180Mbit ≡ 45 books***

→ compression possible to 4.8 kBit/s (1:10)

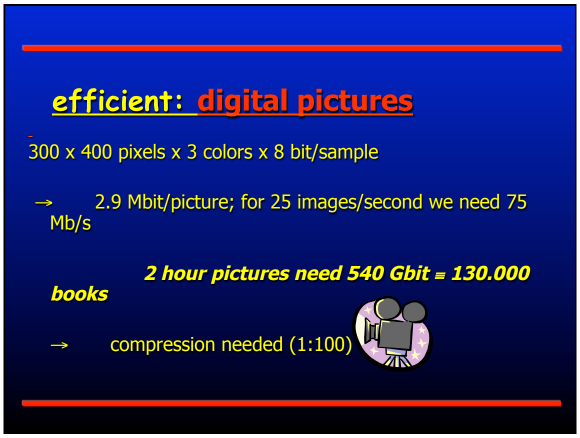


efficient: CD music

sampling speed 44.1 k samples/sec; accuracy 16 bits/sample

→ ***storage capacity for one hour stereo: 5 Gbit ≡ 1250 books***

→ compression possible to 4 bits/sample (1:4)



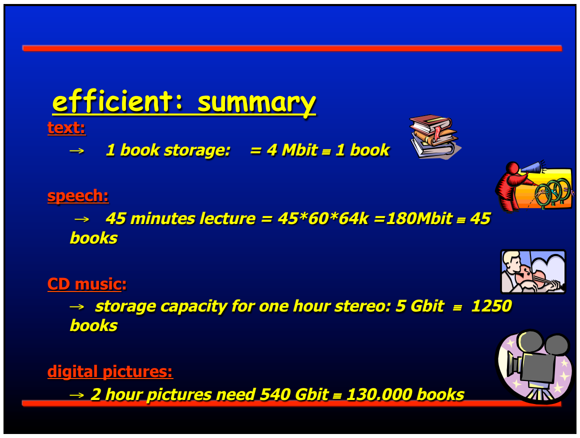
efficient: digital pictures

300 x 400 pixels x 3 colors x 8 bit/sample

→ 2.9 Mbit/picture; for 25 images/second we need 75 Mb/s

2 hour pictures need 540 Gbit ≡ 130.000 books

→ compression needed (1:100)



efficient: summary

text:

→ ***1 book storage: = 4 Mbit ≡ 1 book***

speech:

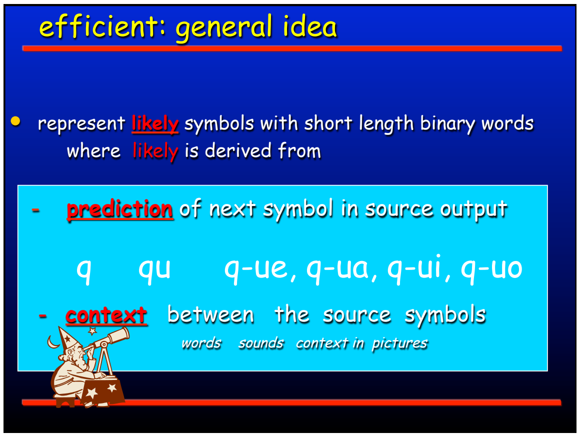
→ ***45 minutes lecture = 45*60*64k =180Mbit ≡ 45 books***

CD music:

→ ***storage capacity for one hour stereo: 5 Gbit ≡ 1250 books***

digital pictures:

→ ***2 hour pictures need 540 Gbit ≡ 130.000 books***



efficient: general idea

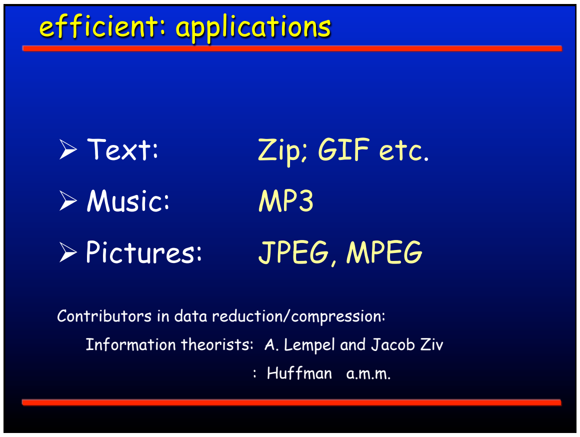
- represent likely symbols with short length binary words where likely is derived from

- prediction of next symbol in source output

q qu q-ue, q-ua, q-ui, q-uo

- context between the source symbols

words sounds context in pictures



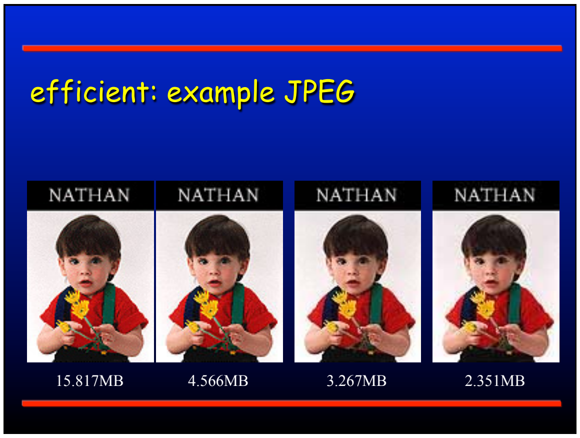
efficient: applications

- Text: Zip; GIF etc.
- Music: MP3
- Pictures: JPEG, MPEG

Contributors in data reduction/compression:

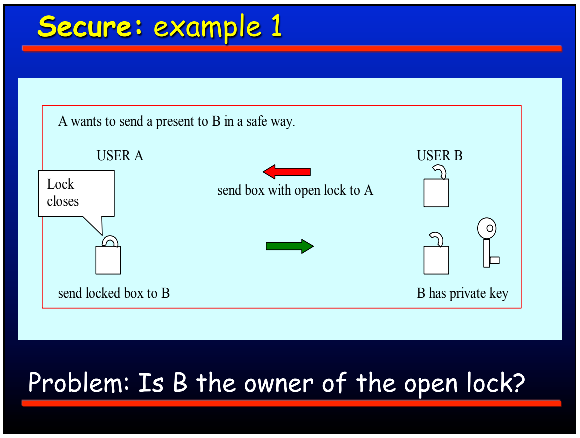
Information theorists: A. Lempel and Jacob Ziv

: Huffman a.m.m.



efficient: example JPEG

NATHAN	NATHAN	NATHAN	NATHAN
15.817MB	4.566MB	3.267MB	2.351MB



Secure: example 1

A wants to send a present to B in a safe way.

USER A

USER B

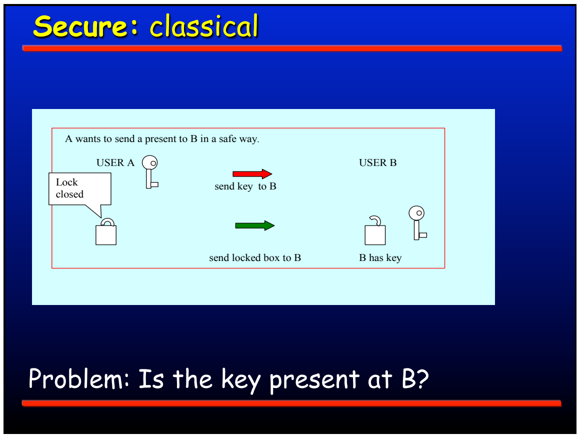
send box with open lock to A

Lock closes

send locked box to B

B has private key

Problem: Is B the owner of the open lock?



Secure: classical

A wants to send a present to B in a safe way.

USER A

USER B

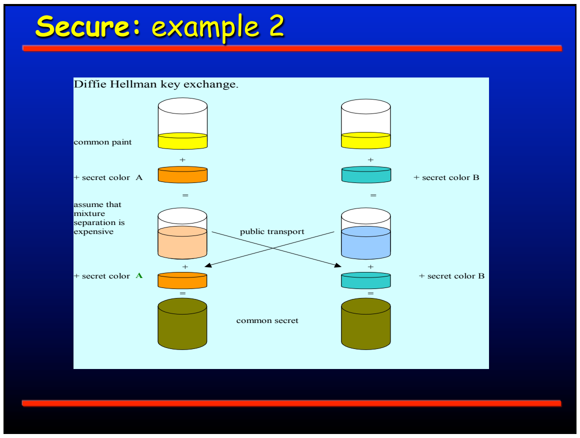
send key to B

Lock closed

send locked box to B

B has key

Problem: Is the key present at B?



Secure: example 2

Diffie Hellman key exchange.

common paint

+ secret color A

+ secret color B

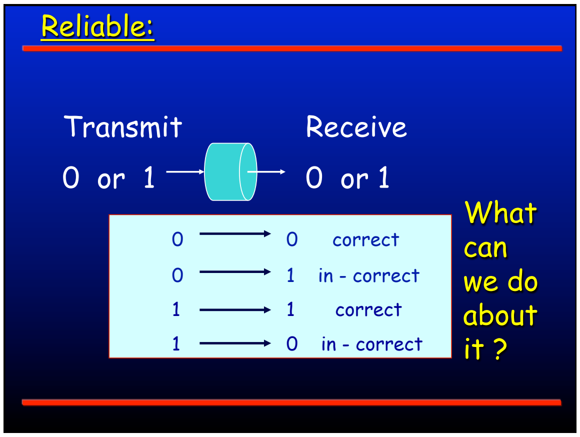
assume that mixture separation is expensive

public transport

+ secret color A

+ secret color B

common secret



Reliable:

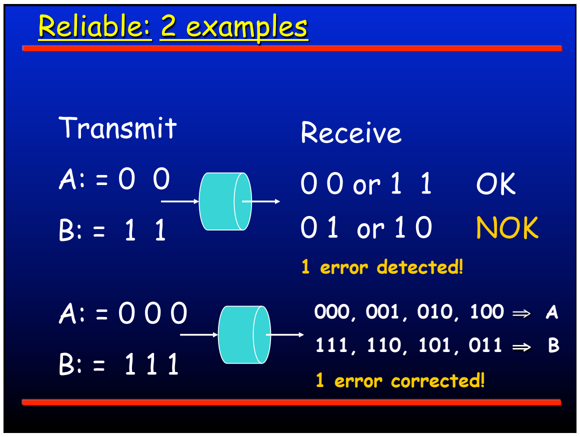
Transmit

Receive

0 or 1 → 0 or 1

0	→	0	correct
0	→	1	in - correct
1	→	1	correct
1	→	0	in - correct

What can we do about it ?



Reliable: 2 examples

Transmit

Receive

A: = 0 0 → 0 0 or 1 1 OK

B: = 1 1 → 0 1 or 1 0 NOK

1 error detected!

A: = 0 0 0

B: = 1 1 1

000, 001, 010, 100 ⇒ A

111, 110, 101, 011 ⇒ B

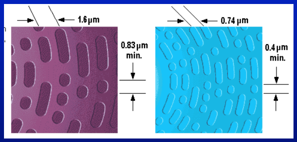
1 error corrected!

Error Sensitivity: Illustration

Error sensitivity: 0.0001=0.01%

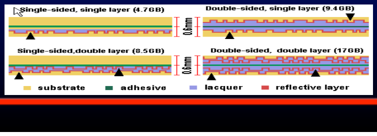
Error sensitivity: 0.0005=0.05%

Optical Storage

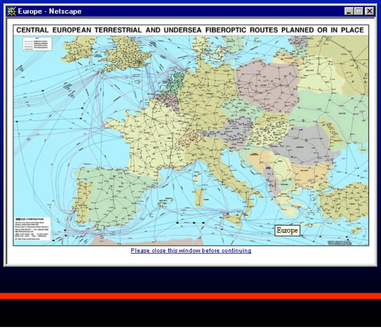


•DVD's seven-fold increase in data capacity over the CD has been largely achieved by tightening up the tolerances throughout the predecessor system

•The data structure was made more efficient by using a better, more efficient error correction code system.


Single-sided, single layer (4.7GB)
Double-sided, single layer (9.4GB)
Single-sided,double layer (8.5GB)
Double-sided, double layer (17GB)
0.6mm
0.6mm
substrate
adhesive
lacquer
reflective layer

Errors in networking



CENTRAL EUROPEAN TERRESTRIAL AND UNDERSEA FIBEROPTIC ROUTES PLANNED OR IN PLACE

Please close this window before continuing