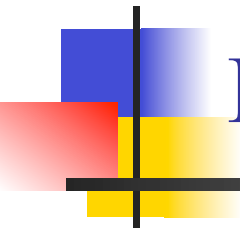
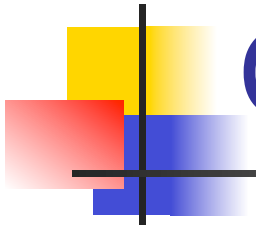


Corso di Crittografia

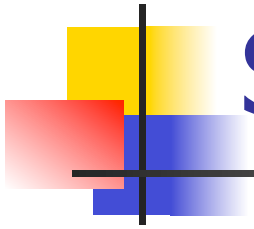
Prof. Dario Catalano

Cifrari Storici e One-Time Pad



Cifrari Storici

- Si considerano tali tutti i cifrari ideati prima degli anni 70
- Sono (quasi) tutti assolutamente insicuri (vedremo perché)
- Qui ne descriveremo (e romperemo) un paio.



Shift Cipher

- K={0,1,…,25}
- Associamo ad ogni lettera un numero
  - A=0, B=1, C=2, …
- Si cifra “spostando” ogni lettera secondo quanto indicato dalla chiave

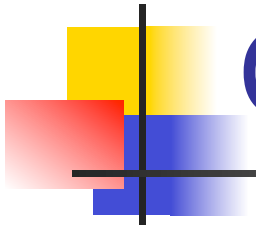
Shift Cipher

Es. k=3, m=dado

dado → 03 00 03 14

"Applichiamo" la chiave

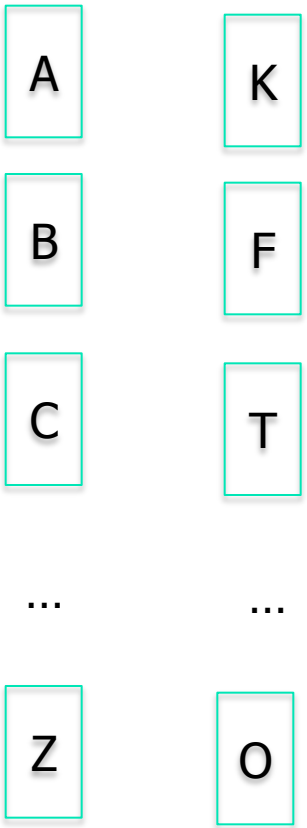
06 03 06 17 → gdgr



Cifrario per sostituzione

- K è l'insieme di tutte le possibili permutazioni dei simboli 0,1,…, 25
- Ogni parola viene cifrata sostituendo al simbolo iniziale, il simbolo indicato dalla chiave

Cifrario per sostituzione: Intuizione



Dunque alla A "sostituiremo" la K, alla B la F, …

Esempio:

k associa: A → K, C → T, S→ D

CASA => TKDK

Semplice analisi

- Il numero di chiavi possibili è molto alto
- Es alfabeto di 256 caratteri
- #chiavi = 256!

$$256! > 128^{128} = 2^{7*128} = 2^{896}$$

Semplice analisi

$$2^{896} > 10^{224}$$

- Supponiamo di avere un computer capace di provare 1000 miliardi di chiavi al secondo
  - 10^{12} chiavi al secondo
- 10^{212} secondi
- > 10^{200} giorni
- > 10^{195} anni
- > 10^{185} miliardi di anni

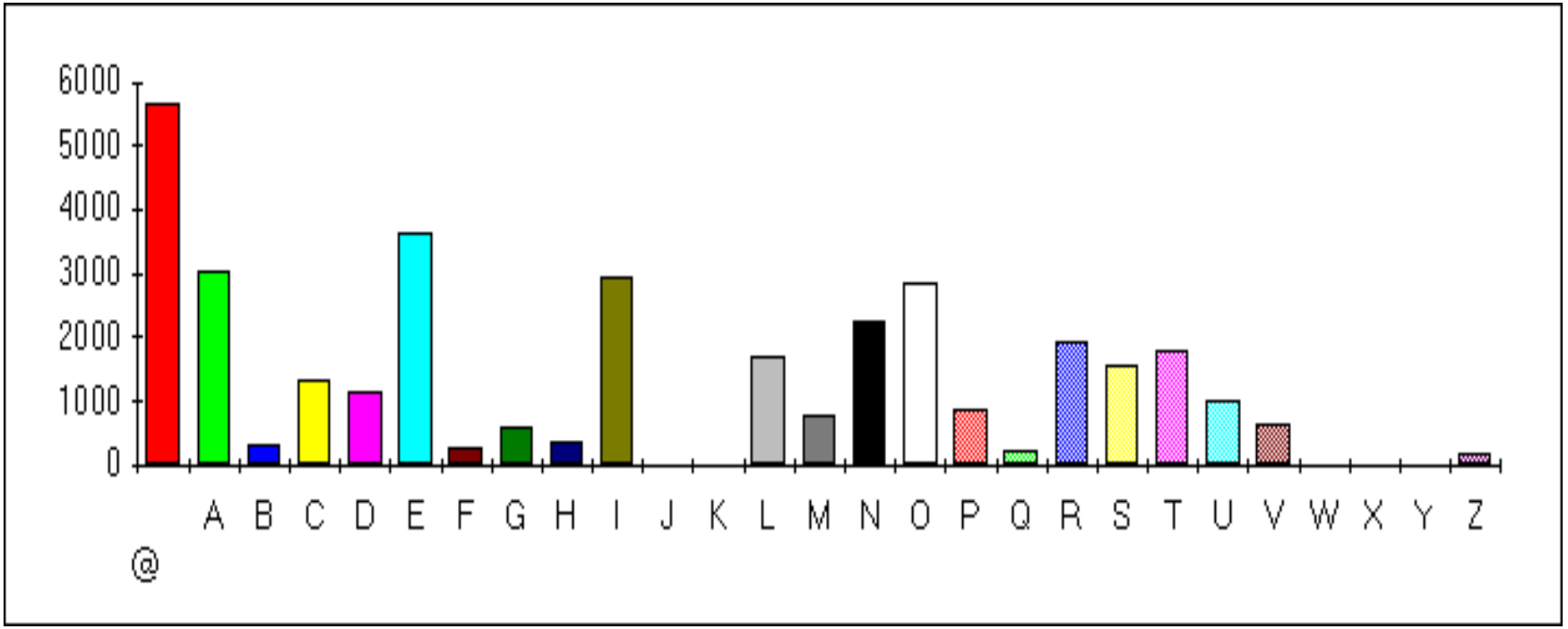
Età dell'universo stimata:
14 miliardi di anni

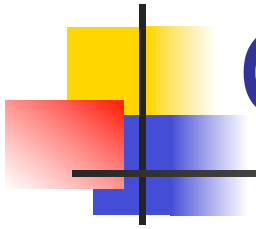
Frequenze della Lingua Italiana (Informazioni tratte da internet)

- Analizzando un testo e calcolando la frequenza delle varie lettere, si può notare che le vocali E, A, O sono le lettere più frequenti.
- Seguono le consonanti L, N, R, S, T .
- Tra i bigrammi più frequenti troviamo Q-U (seguiti da vocale).
- La lettera H é spesso preceduta dalla lettera C o G per formare i trigrammi CHE, CHI, GHE, GHI .
- Nella lingua italiana sono quasi assenti le lettere J, K, Y, X, W.

Un esempio

- Il seguente grafico mostra la frequenza delle varie lettere del I capitolo de "I Promessi Sposi"





Crittanalisi di un testo

lkllm om aztbcobm amobdb efgf ebtbob om kgm zoom zolhm fcgb aztbcobz bgamobdm m bgamobdm z tfif ekf

lkllf mhz efllfefnhz bg dzez lz tfcobm mhz pmgklz z eznmhm dqm bo tzhblf zpmpz kgz hmozrbfgm dfg oz cfpmhgzglm ahzgdmem dqm mhz elzlz nhmeef ib ofhf, m zpmpz ibdqbzhzlf zo tzhblf ib gfg nflmh nbk pbpmhm dfg okb gmooz elmeez dzez

Statistiche del testo in esame

M - 37	E – 15
Z – 36	D – 11
B – 27	P, K – 8
F – 23	T – 7
O – 20	A – 6
L – 19	C,N – 5
G – 17	I – 4
H – 16	Q – 3, R – 1

Cominciamo l'analisi…

- Le lettere che appaiono piu' frequentemente sono M (37), Z (36) e B (27).
- In base alle statistiche a disposizione, ognuna di esse potrebbe cifrare le vocali a, e oppure o

Guardando il testo

lkllm om aztbcobm amobdb efgf ebtbob om kgm zoom zolhm fcgb aztbcobz bgamobdm **m** bgamobdm **z** tfif ekf

lkllf mhz efllfefnhz bg dzez lz tfcobm mhz pmgklz **z** eznmhm dqm bo tzhblf zpmpz kgz hmozrbfgm dfg oz cfpmhgzglm ahzgdmem dqm mhz elzlz nhmeef ib ofhf, **m** zpmpz ibdqbzhzlf zo tzhblf ib gfg nflmh nbk pbpmhm dfg okb gmooz elmeez dzez

- Sia la m che la z appaiono da sole
- Questo suggerisce che possano essere vocali
- Proviamo **z => a**; **m => e**

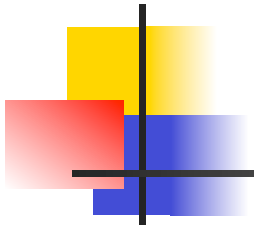
Ecco cosa otteniamo

lkllE oE aAtbcobE aEobdb efgf ebtbob oE kgE AooE AolhE fcgb aAtbcobA bgaEobdE E bgaEobdE A tfif ekf
lkllf EhA efllfefnhA bg dAeA lA tfcobE EhA pEgklA A eAnEhE dqE bo tAhblf ApEpA kgA hEoArbfgE dfg oA cfpEhgAglE ahAgdEeE dqE EhA elAlA nhEeef ib ofhf E ApEpA ibdqbAhAlf Ao tAhblf ib gfg nflEh nbk pbpEhE dfg okb gEooA elEeeA dAeA

Andiamo avanti

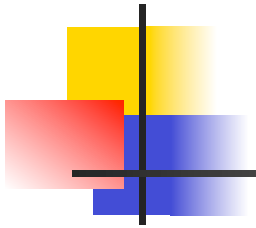
lkllE oE aAtbcobE aEobdb efgf ebtbob oE kgE AooE AolhE fcgb aAtbcobA bgaEobdE E bgaEobdE A tfif ekf
lkllf EhA efllfefnhA bg dAeA lA tfcobE EhA pEgklA A eAnEhE dqE bo tAhblf ApEpA kgA hEoArbfgE dfg oA cfpEhgAglE ahAgdEeE dqE EhA elAlA nhEeef ib ofhf E ApEpA ibdqbAhAlf Ao tAhblf ib gfg nflEh nbk pbpEhE dfg okb gEooA elEeeA dAeA

- La b è molto frequente (27)
- Inoltre appare spesso come lettera finale
- Analogo discorso per la f (23)
- E' plausibile che esse codifichino vocali.
  - Tentiamo **b=>i f=>o**



lkllE oE aAtIcoIE aEoIdI eOgO eItIoI oE kgE AooE AolhE OcgI
aAtIcoIA IgaEoIdE E IgaEoIdE A tOiO ekO
lkllO EhA eOllOeOnhA Ig dAeA lA tOcoIE EhA pEgklA A eAnEhE dqE
Io tAhIlO ApEpA kgA hEoArIOgE dOg oA cOpEhgAglE ahAgdEeE
dqE EhA elAlA nhEeeO iI oOhO, E ApEpA iIdqIAhAlO Ao tAhIlO iI
gOg nOlEh nIk pIpEhE dOg okI gEooA elEeeA dAeA

- Non abbiamo ancora inserito consonanti. Tra le consonanti più frequenti in italiano vi sono la N e la L.
- La lettera g appare 17 volte.
  - Inoltre essa è presente sia all'inizio che alla fine di parole, potrebbe codficare la N
- Per analoghe ragioni la o potrebbe codificare la L

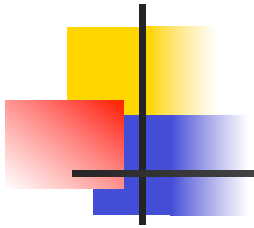


lkllE LE aAtIcLIE aELIdI eONO eItILI LE kNE ALLE ALlhE OcNI aAtIcLIA INaELIdE E INaELIdE A tOiO ekO

lkllO EhA eOllOeOnhA IN dAeA LA tOcLIE EhA pENklA A eAnEhE dqE IL

tAhIlO ApEpA kNA hELArIONE dON LA cOpEhNANlE ahANdEeE dqE EhA elAlA nhEeeO iI LOhO, E ApEpA iIdqIAhAlO AL tAhIlO iI NON nOlEh nIk pIpEhE dON LkI NELLA elEeeA dAeA

- In base al testo la h (16) potrebbe codificare la R (lettera molto frequente)
- Analogamente, la p potrebbe codificare la V

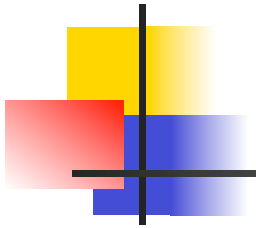


lkllE LE aAtIcLIE aELIdI eONO eItILI LE kNE ALLE ALlRE OcNI aAtIcLIA INaELIdE E INaELIdE A tOiO ekO
lkllO ERA eOllOeOnRA IN dAeA LA tOcLIE ERA VENklA A eAnERE dqE IL tARIlO AVEVA kNA RELArIONE dON LA cOVERNANlE aRANdEeE dqE ERA elAlA nREeeO iI LORO, E AVEVA iIdqIARAlO AL tARIlO iI NON nOlER nIk VIVERE dON LkI NELLA elEeeA dAeA

- In base al testo la r certamente codifica la Z
- Analogamente, la d dovrebbe codificare la C e la l la T

TkTTE LE aAtIcLIE aELICI eONO eItILI LE kNE ALLE ALTRE OcNI
aAtIcLIA INaELICE E INaELICE A tOiO ekO
TkTTO ERA eOTTOeOnRA IN CAeA LA tOcLIE ERA VENkTA A eAnERE
CqE IL tARITO AVEVA kNA RELAZIONE CON LA cOVERNANTE aRANCEeE
CqE ERA eTATA nREeeO iI LORO, E AVEVA iICqIARATO AL tARITO iI NON
nOTER nIk VIVERE CON LkI NELLA eTEeeA CAeA

- Ovvie sostituzioni:
  - e=> S
  - k=> U
  - a=> F
  - t=> M



TUTTE LE FAMIcLIE FELICI SONO SIMILI LE UNE ALLE ALTRE OcNI FAMIcLIA INFELICE E INFELICE A MOiO SUO
TUTTO ERA SOTTOSOnRA IN CASA LA MOcLIE ERA VENUTA A SAnERE CqE IL MARITO AVEVA UNA RELAZIONE CON LA cOVERNANTE FRANCESE CqE ERA STATA nRESSO iI LORO, E AVEVA iICqIARATO AL MARITO iI NON nOTER nIU VIVERE CON LUI NELLA STESSA CASA

Il resto possiamo indovinarlo facilmente…

Tutte le famiglie felici sono simili le une alle altre; ogni famiglia infelice è infelice a modo suo.

Tutto era sottosopra in casa. La moglie era venuta a sapere che il marito aveva una relazione con la governante francese che era stata presso di loro, e aveva dichiarato al marito di non poter più vivere con lui nella stessa casa.

L.Tolstoj "Anna Karenina"

Limitazioni della Perfetta Sicurezza

Teorema

Sia (KeyGen, Enc, Dec) uno schema perfettamente sicuro ($\mathcal{M}$ spazio dei messaggi, $\mathcal{K}$ spazio delle chiavi) allora

$$|\mathcal{K}| \geq |\mathcal{M}|$$

Caratterizzazione della perfetta sicurezza

Teorema (Shannon)

SE=(KeyGen, Enc, Dec), $|\mathcal{M}|=|\mathcal{K}|=|C|$ SE offre perfetta sicurezza se e soltanto se

1. $\forall k \in \mathcal{K}$, k è scelta con probabilità $1/|\mathcal{K}|$
2. $\forall m \in \mathcal{M}, \forall c \in C, \exists ! k \in \mathcal{K}$, tale che

$$\mathrm{Enc}(k,m)=c$$