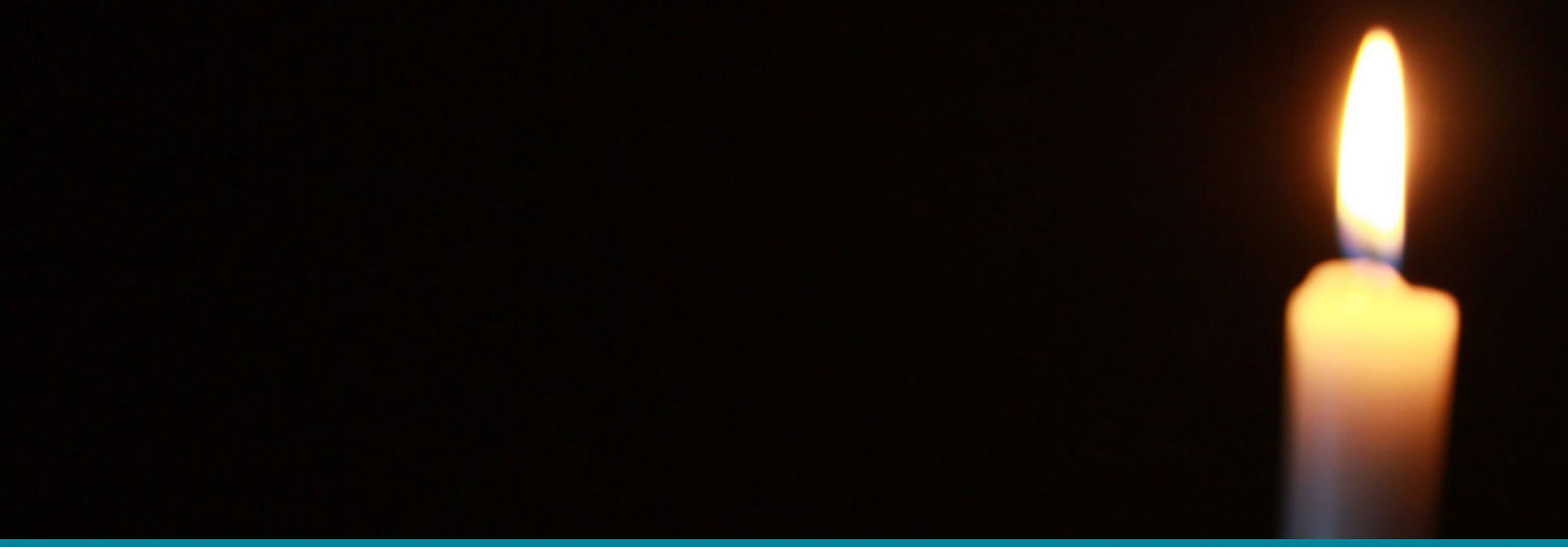


Crittografia - 1

L'arte di svelare segreti!

RESEARCH IN ACTION - RIA

RESEARCHINACTION.IT

Il desiderio di svelare segreti è profondamente radicato nella natura umana; la promessa di partecipare a conoscenze negate ad altri eccita anche la mente meno curiosa. Qualcuno ha la fortuna di trovare un lavoro che consiste nella soluzione di misteri, ma la maggior parte di noi è spinta a soddisfare questo desiderio risolvendo enigmi artificiali ideati per il nostro divertimento. I romanzi polizieschi o i cruciverba sono rivolti alla maggioranza; la soluzione di codici segreti può essere l'occupazione di pochi.

John Chadwick, *Lineare B: l'enigma della scrittura micenea.*

13 Crittografia 1 - 07.19
Revisione 0.1 14.11.19

RiA - Research in Action

La parola ria in inglese significa estuario, in particolare (dalla definizione che ne da l'Oxford Living Dictonaries):

> A long, narrow inlet formed by the partial submergence of a river valley ... the rias or estuaries contain very peculiar ecosystems which often contain important amounts of fish ... (a causa della loro natura, le rias o estuari contengono ecosistemi molto particolari che spesso contengono grandi quantità di pesce - www.eurotomic.com/spain/the-rias-altas-in-spain.php)

quindi questo prodotto che sarà realizzato grazie all'attività di alternanza scuola-lavoro di alcuni studenti del liceo scientifico G.B.Grassi di Latina - **www.liceograssilatina.org** - sarà un luogo virtuale da esplorare dove *pescare* molto materiale per la didattica laboratoriale.

Fare scienza

La scienza non è solo identificabile con la formula, il modello, la teoria. In altre parole la scienza non rappresenta solo un corpo di conoscenze organizzate e formalizzate. La scienza è anche e fondamentalmente ricerca. Una ricerca volta a conoscere e a capire sempre più e sempre meglio come è fatto e come funziona questo nostro complicatissimo mondo.

Fare scienza si identifica con l'interrogarsi, con l'indagare ed esplorare fatti e cose. Questo tipo di lavoro i bambini lo fanno spontaneamente sin dalla loro nascita ma si perde nel corso del percorso scolastico. L'intervento educativo deve tener conto di ciò e fornire stimoli, occasioni e strumenti per far acquisire agli studenti capacità sempre più ampie e affinate per poter compiere questo lavoro di indagine mantenendo viva (o risvegliando) la curiosità cognitiva, la voglia di sapere e di scoprire, la fiducia di poter capire.

Pensare in senso creativo, in campo scientifico, significa aggredire i problemi, attivare processi vivi dei pensiero, alimentare l'evoluzione dinamica dell'intelligenza duttile, dell'esercizio dell'intuizione e dell'immaginazione, della capacità di progettare e formulare ipotesi, di controllare e verificare quanto prodotto e ricercato.

Per questo è necessario bandire forme di apprendimento consumate entro schemi rigidi di elaborazione del pensiero e puntare al recupero della congettura, dell'ipotesi, di una coscienza scientifica aperta a interrogare ogni problematica.

La società odierna deve far fronte ad un rinnovamento scientifico e tecnico accelerato in cui lo sviluppo delle conoscenze scientifiche e la creazione di prodotti di alta tecnologia *(hi-tech)*, come anche la loro diffusione subiscono un'accelerazione sempre più rapida.

È necessaria, quindi, una diffusione della conoscenza in genere ed è indispensabile promuovere una nuova cultura scientifica e tecnica basata sull'informazione e sulla conoscenza. E quanto più è solida la base di conoscenze scientifiche scolastiche, tanto più si può approfittare dell'informazione e della conoscenza scientifica e tecnica.

» https://www.facebook.com/Research-in-Action-341307966417448/
» https://www.youtube.com/channel/UC1PA7Zu78RUMBJnkaiOR8kA/

Sommario dei contenuti

Crittografia 1 - L'arte di svelare segreti!

Sommario dei contenuti

Blockly è un generatore di codice open source sostenuto da Google: https://developers.google.com/blockly/.

Materiale disponibile per questo laboratorio:

» il fascicolo con la prima parte (questa) del laboratorio (in formato PDF di circa 12MB): **http://researchinaction.it/materials/13-Crittografia-1.pdf**;

» il fascicolo con la seconda parte del laboratorio (in formato PDF di circa 6MB): **http://researchinaction.it/materials/14-Crittografia-2.pdf**;

» il simulatore di Enigma è descritto qui: **http://researchinaction.it/i-nostri-giochi/simulatore-enigma/**;

» la libreria, sviluppata con Blockly, si può provare sul campo: **http://researchinaction.it/myblockly/criptography.html** ed è descritta nel dettaglio qui: **http://researchinaction.it/laboratori/crittografia/**.

Per il materiale didattico a supporto del fascicolo visitare anche la pagina Download del sito dedicato al progetto: **http://researchinaction.it/download/**.

Per i videotutorial è possibile visitare il canale YouTube del progetto: **https://www.youtube.com/channel/UC1PA7Zu78RUMBJnkaiOR8kA**.

Crittografia 1

L'arte di svelare segreti!

1. Introduzione

Questo laboratorio è stato sviluppato per la prima classe del liceo matematico ed è nato da una conversazione con la professoressa Lucilla Galterio, il progetto è stato coordinato da Gualtiero Grassucci (Iss G.B. Grassi di Latina).

Il desiderio di svelare segreti è profondamente radicato nella natura umana; la promessa di partecipare a conoscenze negate ad altri eccita anche la mente meno curiosa. Qualcuno ha la fortuna di trovare un lavoro che consiste nella soluzione di misteri, ma la maggior parte di noi è spinta a soddisfare questo desiderio risolvendo enigmi artificiali ideati per il nostro divertimento. I romanzi polizieschi o i cruciverba sono rivolti alla maggioranza; la soluzione di codici segreti può essere l'occupazione di pochi.

John Chadwick, Lineare B: l'enigma della scrittura micenea.

Per approfondire gli argomenti affrontati in questo laboratorio è possibile leggere, tra gli altri, Singh, S. - Codici & segreti. La storia affascinante dei messaggi cifrati dall'antico Egitto a internet.

La crittografia è molto antica. Fin dall'inizio dei tempi infatti si è cercato di celare il significato di messaggi e lettere o di rendere oscuro il significato di comunicazioni o dispacci in modo che solo il destinatario potesse interpretare i simboli e chiarire il significato del testo. D'altra parte, accanto a questo bisogno di segretezza, si sono sviluppate tecniche e metodi per carpire i segreti nascosti in quei messaggi. Tecniche evolute, nel corso dei secoli, grazie al contributo di personaggi tra i più disparati e insospettabili come Leon Battista Alberti o, prima di lui, Gaio Giulio Cesare, autore di un metoso che, per quanto semplice, ancora oggi porta il suo nome.

Oggi più che mai siamo a contatto quotidianamente, anche se praticamente a nostra insaputa, con tecniche di cifratura e protezione di dati, informazioni e messaggi, da quando inviamo un messaggio dal nostro *smartphone* a quando usiamo la carta di credito per un acquisto *online*.

1.1. Alcune definizioni

L'obiettivo della cifratura è quello di garantire la sicurezza di un messaggio, sicurezza che avviene cifrando il testo attraverso un sistema che permetta, in qualche modo, di sostituire all'originale, detto testo in chiaro, un testo cifrato che potrebbe anche non usare lo stesso insieme di simboli. Comunemente il testo, e l'alfabeto, in chiaro sono scritti con lettere minuscole mentre il testo cifrato usa solo lettere maiuscole, convenzione che seguiremo anche in questo laboratorio.

Il mittente è colui che, una volta cifrato il testo, invia il messaggio al destinatario. Quest'ultimo provvede all'interpretazione del messaggio per trasformare il testo cifrato nuovamente in chiaro.

Anche se spesso usati come sinonimi i verbi cifrare e codificare hanno un significato leggermente diverso:

» con cifrare si intende l'operazione di nascondere il significato di un messaggio sostituendo alle lettere del testo in chiaro altre lettere o simboli;
» con codifcare invece si raggiunge lo stesso scopo sostituendo intere parole o frasi.

La chiave è l'informazione usata per l'operazione di cifratura o di codifica e, in qualche modo, deve essere conosciuta anche dal destinatario perché possa ricostruire il messaggio originale (decifrare o decodificare).

Molto spesso l'alfabeto usato per scrivere un messaggio leggibile, non cifrato, è detto alfabeto in

chiaro e, allo stesso modo, il testo è detto messaggio in chiaro; alfabeto e messaggio in chiaro sono usualmente in caratteri minuscoli. L'insieme dei simboli usati per cifrare il testo è detto alfabeto cifrato e il testo criptato è chiamato messaggio/testo cifrato; di consueto per l'alfabeto e per il messaggio cifrato si usano caratteri maiuscoli.

Gli obiettivi della crittografia sono fondamentalmente due e sostanzialmente in concorrenza tra loro: da una parte si cerca di sviluppare tecniche sempre più sofisticate per proteggere le informazioni e impedire accessi indesiderati ai dati, dall'altra si cerca di evolvere meccanismi per forzare tali protezioni e raggiungere quelle informazioni che con tanta cura si sono nascoste. I due aspetti, nei secoli, sono andati di pari passo e a ogni progresso in un campo quasi sempre ha fatto seguito la nascita di una tecnica o di una tecnologia per contrastare quel progresso.

In questo laboratorio affronteremo la prima parte di un percorso volto alla conoscenza e sperimentazione delle tecniche di crittografia e criptoanalisi più famose e importanti dal punto di vista storico, anche se spesso ormai obsolete.

1.2. Prerequisiti

Per affrontare questo laboratorio sono necessarie:

» conoscenze elementari di calcolo combinatorio (sostanzialmente disposizioni semplici, combinazioni semplici, permutazioni e la capacità di risolvere semplici problemi);
» conoscenze elementari dell'artimetica modulare (essenzialmente la capacità di calcolare un'espressione del tipo $k \bmod n$)
» la conoscenza dei concetti elementari relativi alle percentuali e alle tabelle a doppia entrata)
» possono essere di aiuto (ma non sono essenziali) anche elementari conoscenze dei concetti di algoritmo, procedura, *coding*;
» la capacità di usare, anche in modo elementare, un'applicazione per gestire testi (come Microsoft Word o Google Documenti), pur non essendo indispensabile, può tornare utile **(cfr. Un suggerimento a pagina 20)**.

1.3. Obiettivi

Gli obiettivi della prima parte sono essenzialmente:

» la conoscenza di alcuni metodi di crittografia;
» l'applicazione di elementari tecniche di criptoanalisi;
» la capacità di valutare, per via empirica, la sicurezza di un metodo di cifratura.

2. Il metodo di Cesare

Con ogni probabilità, il metodo cifratura da cui abbiamo scelto di iniziare, è il primo di cui abbiamo documentazione che sia stato pensato e progettato specificatamente per uso militare ed è stato, spesso attribuito al generale romano, e poi dittatore, Gaio Giulio Cesare (100 a.C. - 44 a.C.).

Sul blog Research in Action è possibile utilizzare la libreria sviluppata con Blockly nel corso delle attività del progetto: http://researchinaction.it/myblockly/criptography.html con cui provare a usare i metodi di cifratura e di analisi affrontati. Blockly è un generatore di codice open source sostenuto da Google: https://developers.google.com/blockly/.

L'algoritmo prevede di sostituire a ciascuna lettera dell'alfabeto un'altra che si trova un certo numero di posti *più in là* nell'alfabeto. Si tratta quindi di un metodo di cifratura perchè opera sulle lettere e non su parole o frasi **(cfr. 1.1 Alcune definizioni a pagina 5)** ed è un algoritmo di sostituzione (a ogni carattere del messaggio in chiaro è sostituito un simbolo dell'alfabeto cifrato). Lo spostamento (... *più in là* ...), un numero intero, è la *chiave*.

Per esempio, se scegliamo come chiave il numero 5 e facciamo riferimento all'alfabeto italiano di 21 lettere, la frase *il desiderio di svelare segreti* diventa *PQ ILAPILZPT IP ADLQFZL ALNZLBP*. Per ottenere il testo cifrato abbiamo sostituito ogni lettera della frase con la lettera che si trova cinque posti avanti nell'alfabeto italiano, per cui al posto della lettera *i* si trova la *P*, al posto della lettera *l* la lettera *Q*. Notare che l'alfabeto è considerato *circolare* per cui alla lettera *v* si sostituisce la lettera *D* perché dopo aver raggiunto la zeta si prosegue contando dalla lettera *a*.

È importante anche notare anche che abbiamo dovuto specificare l'alfabeto, infatti la stessa frase, con la stessa chiave, usando l'alfabeto inglese sarebbe diventata *NQ IJXNIJWNT IN XAJQFWJ XJLWJYN*.

La decifratura del messaggio, conoscendo la chiave usata per la cifratura, è piuttosto semplice: è sufficiente sostituire a ogni lettera quella che si trova un certo numero di posti *prima* nell'alfabeto. Nel caso della frase cifrata *NQ IJXNIJWNT IN XAJQFWJ XJLWJYN*, ricordando che la chiave è 5, la decifratura ci dà: *il desiderio di svelare segreti*. Infatti al simbolo *N* abbiamo sostituito la lettera che, nell'alfabeto, la precede di cinque posti, la *i*; al simbolo *Q* la lettera *l*, e così via ...

2.1. Alcuni suggerimenti

Per facilitare la cifratura e la decifratura è utile costruire una tabella di traduzione con l'alfabeto in chiaro e quello cifrato su due righe in modo da avere il carattere cifrato in corrispondenza del corrispondente in chiaro. Questa che segue è la tabella per il metodo di Cesare con chiave pari a cinque (la stessa usata nell'esempio che abbiamo appena fatto).

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
F	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

Per cifrare un messaggio si cerca il carattere nella prima riga e si sostituisce con quello corrispondente (nella stessa colonna) che si trova nella seconda riga. Per decifrare, al contrario, si cerca il carattere nella seconda riga e si sostituisce con quello che si trova nella stessa colonna ma nella prima riga. Se ora si prova di nuovo a cifrare la frase *il desiderio di svelare segreti* si vede subito che il procedimento è reso più veloce dalla tabella. Questa tabella risulterà utile in ogni algoritmo cifratura mediante sostituzione.

Se si lavora a cifratura e decifratura usando carta e penna si deve cercare di distinguere i caratteri in chiaro (di solito minuscoli) da quelli cifrati (in genere maiuscoli), come per esempio la *l* (elle) minuscola e la *I* maiuscola. Allo stesso modo, se usate un dispositivo elettronico si devono scegliere caratteri *con le grazie* come questo usato per il testo che state leggendo e non caratteri *senza grazie* come questo per lo stesso motivo.

Può essere conveniente, in modo particolare se il testo è piuttosto lungo, sostituire tutte le occorrenze della stessa lettera piuttosto che sostituire la prima, la seconda, … Per esempio, facendo riferimento alla frase *il desiderio di svelare segreti* conviene sostituire tutte le *i*, poi tutte le *l* (elle) e così via.

2.2. Precisazioni

Ripulitura del testo prima della cifratura

Per complicare un poco di più la vita a chi tenta di svelare il messaggio senza averne l'autorizzazione, è utile eliminare spazi, segni di punteggiatura e altri caratteri che non siano alfanumerici per cui la nostra frase, prima della cifratura ma dopo la fase di preparazione, sarebbe stata *ildesideriodisvelaresegreti* e una volta cifrata sarebbe diventata *NQIJXNIJWNTINXAJQFWJXJLWJYN*.

Operatore modulo

Dal punto di vista matematico, il metodo di Cesare con chiave pari a c (dove c è un numero intero) sostituisce alla lettera che si trova in posizione n (dove n è compreso tra 1 e la lunghezza dell'alfabeto, 26 per quello anglosassone) il carattere dell'alfabeto cifrato che si trova in posizione $n + c$ ma, com'è ovvio, $n + c$ non può superare 26 per cui si ricorre all'operatore modulo che restituisce il resto della divisione di $n + c$ per la lunghezza dell'alfabeto, se usiamo l'alfabeto anglosassone, la posizione del simbolo cifrato si calcola come

$$(n + c) \bmod 26$$

il resto della divisione per 26 è sempre compreso tra 0 e 25 e quindi la sostituzione è possibile.

Per esempio, la posizione della lettera *w* nell'alfabeto anglosassone è $n = 22$ (iniziando a contare da zero), se la chiave è $c = 5$ (come negli esempi) la posizione del simbolo cifrato dovrebbe essere $n + c = 22 + 5 = 27$: impossibile visto che l'alfabeto cifrato contiene solo 26 simboli! In precedenza abbiamo risolto contando fino a 26 e poi ricominciando dall'inizio ma se proviamo a usare la formula che abbiamo dato qui sopra otteniamo

$$(22 + 5) \bmod 26 = 1$$

infatti il resto della divisione di 27 per 26 è proprio 1. Il simbolo da sostituire è quello che si trova nella posizione uno dell'alfabeto cifrato: la *B* (considerando che la *A* è in posizione zero).

Il modulo di due numeri interi *a* e *b* è *a mod b* ed è definito come il resto della divisione di *a* per *b*. Siamo in un certo senso abituati all'operatore modulo perchè lo usiamo quotidianamente nel calcolo delle ore: se abbiamo una riunione alle 11 di mattina che durerà 3 ore sappiamo che finirà alle 2 del pomeriggio, infatti (11+3) mod 12 è 2 (il resto della divisione di 14 per 12. Per esempio: 7 mod 26 è 7 perchè dividendo 26 per 7 otteniamo zero come quoziente e 7 come resto, ma anche 33 mod 26 fa 7 perchè 33 diviso 26 fa uno con il resto ancora di 7 mentre 26 mod 26 fa zero!
È bene notare che sia *a* che *b* possono essere numeri negativi per cui, per esempio, -3 mod 26 fa 23, infatti il resto della divisione di -3 per 26 è proprio 23.

2.3. Proviamo

È ora di prendere carta e penna (o un'applicazione per gestire testi) e mettersi al lavoro.

Scegli una chiave (in sostanza un numero intero) e prova a cifrare e decifrare la frase che segue: *Due cose sono infinite: l'universo e la stupidità umana, ma riguardo l'universo ho ancora dei dubbi.* **Ricorda di eliminare tutti i caratteri che non sono alfanumerici prima della cifratura.**

Per evitare errori conviene costruire una tabella come mostrato in precedenza, puoi usare la tabella che trovi qui di seguito. Si tratta di inserire il simbolo corrispondente alla lettera *a* nella prima casella della seconda riga (quello che si trova *c* posti dopo la lettera *a* nell'alfabeto, dove *c* è l'intero scelto come chiave), arrivati alla zeta si prosegue re-iniziando dalla *A*.

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

Come avrai notato l'operazione di decifratura non è affatto difficile conoscendo la chiave. Ma ora immagina di voler decifrare un messaggio che non è destinato a te e quindi non conosci la chiave, le cose non sono così semplici.

Chiedi a qualcuno di cifrare un messaggio a sua scelta e cerca di decifrarlo senza conoscerne la chiave. Quante possibili chiavi ha potuto scegliere il creatore del messaggio cifrato?

Siamo appena all'inizio del laboratorio e quindi più interessati ai principi generali, per ora, e non tanto a questo specifico metodo, per cui è meglio lavorare su messaggi non troppo lunghi.

Sei riuscito a decifrare il messaggio segreto? Come valuti la sicurezza del codice di Cesare?

L'operazione di analisi e decifratura di un messaggio cifrato, senza conoscere la chiave di cifratura, è detta criptoanalisi. Prima di passare ad altro, un'ultima domanda …

Cosa è necessario comunicare al destinatario, oltre al messaggio stesso naturalmente, perché riesca a decifrare correttamente e senza sforzo il testo cifrato.

Si inizia a capire che il testo cifrato non è l'unica cosa che il destinatario deve ricevere perchè possa comprendere il messaggio!

Research in Action
math inside

In matematica quando si tenta di risolvere un problema esplorando tutte le possibili soluzioni si dice di aver risolto il problema usando *la forza bruta*. In alcuni casi questo approccio è facile, efficace e spesso veloce, specie se i tentativi sono automatizzati. In altri è davvero difficile, per il grandissimo numero di soluzioni possibili, o del tutto impraticabile.

3. Cifratura per sostituzione

Ti sarai accorto che, ache senza dispositivi elettronici, è piuttosto facile venire a capo della cifratura con il metodo attribuito a Giulio Cesare. In realtà si può fare meglio con una piccolissima modifica, dove con ... *meglio* ... intendiamo *realizzare un codice più sicuro.*

Nel codice di Cesare la chiave è un numero intero che, per l'alfabeto inglese, può andare da 1 a 25 (in caso la chiave sia uguale a 26 otteniamo come alfabeto cifrato esattamente l'alfabeto in chiaro) ed è proprio questa la debolezza principale di questo algoritmo. Infatti è possibile, in un tempo ragionevole, provare tutte le possibili chiavi fino a individuare quella giusta.

Se invece costruiamo la tabella di cifratura scegliendo a caso le posizioni dei simboli dell'alfabeto cifrato abbiamo una chiave molto più lunga: l'intero alfabeto cifrato, ma anche molte più possibilità e quindi maggiore sicurezza.

In concreto, prendiamo la tabella che abbiamo usato per la cifratura e nella prima riga, come in precedenza, inseriamo l'alfabeto in chiaro. Nella seconda riga, però, mettiamo le 26 lettere dell'alfabeto inglese in modo casuale, come in questo esempio:

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

La cifratura avviene nello stesso modo (si cerca la lettera in chiaro nella prima riga e si sostituisce con la corrispondente nella seconda riga) così come la decifratura, a patto di conoscere la chiave che in questo caso è *JTZKCMLBDNAYUOEVXFPQWGRHIS*. In pratica, come detto, l'intero alfabeto cifrato!

3.1. Proviamo

Scegli una chiave. In altre parole, costruisci una tabella disponendo casualmente i simboli della seconda riga.

Prova a cifrare e decifrare la frase *Due cose sono infinite: l'universo e la stupidità umana, ma riguardo l'universo ho ancora dei dubbi* **con il metodo che abbiamo appena appreso. Ricorda sempre di eliminare tutti i caratteri che non sono alfanumerici prima della cifratura.**

Anche in questo caso puoi usare la tabella che trovi qui di seguito per facilitare il lavoro.

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

Come avrai notato l'operazione di cifratura non è molto più complessa rispetto al metodo di Cesare.

Ora, come in precedenza, mettiamoci nei panni di una spia e proviamo a decifrare il messaggio senza conoscere la chiave, le cose non sono così semplici anzi, diventano veramente complesse.

Chiedi a qualcuno di cifrare un messaggio con una chiave a sua scelta, realizzata come appena descritto, e cerca di decifrarlo senza conoscerne la chiave. Quante possibili chiavi ha potuto scegliere il creatore del messaggio cifrato?

Se tutto è andato bene, dovresti aver capito che l'analisi del testo cifrato è diventata ben più complessa!

Sei riuscito? Come valuti la sicurezza di questo algoritmo di cifratura rispetto al precedente? Può garantire la sicurezza?

Nota a margine: il sito *Parole con* https://www.parolecon.it/ può essere di aiuto: permette infatti di cercare le parole italiane che iniziano, finiscono o contengono una determinata stringa di caratteri.

Come in precedenza, ricorda che per adesso non siamo interessati tanto a questo specifico metodo quanto piuttosto a comprendere i meccanismi della crittografia, quindi cerca di comprenderne vantaggi e svantaggi, in particolare ...

Cosa è necessario comunicare al destinatario, oltre al messaggio stesso naturalmente, perché riesca a decifrare correttamente e senza sforzo il testo cifrato.

E ora mettiamoci nei panni di chi vuole decifrare il messaggio una volta intercettato …

3.2. La criptoanalisi

La criptoanalisi, come accennato all'inizio di questo laboratorio, è l'insieme delle tecniche scoperte nei secoli per decifrare un messaggio senza conoscerne la chiave (e a volte senza nemmeno conoscere l'algoritmo usato per la cifratura). Qui cerchiamo di affrontare una di queste tecniche.

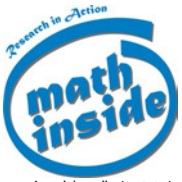
La cifratura per sostituzione, anche questa appena utilizzata, ha un difetto: a ogni carattere dell'alfabeto in chiaro è sostituito sempre lo stesso carattere dell'alfabeto cifrato (basta tornare indietro e riflettere sugli esempi che abbiamo proposto). Questo vuol dire che il numero delle occorrenze di una certa lettera (dell'alfabeto in chiaro) sarà uguale al numero delle occorrenze del simbolo che l'ha sostituita nel testo cifrato.

Quindi potremmo cercare di riconoscere a quale lettera corrisponde un simbolo in un testo cifrato confrontando la frequenza con cui appare nel messaggio rispetto quella con cui appare nella lingua del messaggio. In altre parole, se la lettera *a* in italiano rappresenta circa il 10% del totale e nel messaggio cifrato un certo simbolo appare più o meno il 10% delle volte allora potremmo pensare che si tratti del simbolo corrispondente alla lettera *a*. Questa è l'idea fondamentale!

In italiano, la frequenza delle singole lettere è mostrata nella tabella che segue. Ti sarà utile nel seguito.

Tabella delle frequenze in italiano

	a	b	c	d	e	f	g	h	i	j	k	l	m
%	10.08	0.85	4.03	4.04	12.86	0.98	1.68	1.04	10.99	0.02	0.01	6.46	2.86
	n	o	p	q	r	s	t	u	v	w	x	y	z
%	7.05	9.18	3.07	0.76	6.62	5.60	6.30	3.39	1.23	0.00	0.01	0.02	0.75



In realtà, quella riportata in questa pagina è solo una delle possibili tabelle delle frequenze che si possono realizzare analizzando testi in italiano. Infatti il linguaggio cambia a seconda delle sue applicazioni e un testo, in italiano come in qualsiasi altra lingua, sarà diverso se descrive, che so, un teorema di matematica o racconta una storia d'amore. E cambiando il linguaggio cambiamo anche le frequenze con cui appaiono le singole lettere.
Dal un punto di vista del criptoanalista sarebbe molto più utile conoscere le frequenze dei caratteri calcolate su testi simili a quello poi da decifrare.

Un lavoro da agente segreto

Abbiamo cifrato, con il metodo di sostituzione, un testo piuttosto lungo, originariamente in italiano, usando l'alfabeto anglosassone di 26 lettere. Il risultato è quello che segue:

```
MDLYDKDLEOKEFKDFEBJOMFJQCYYDUDCDGCKEOCDGEPQFDEZZBDYJPQCPPJVJWFJZB-
CVEQFCTTCJMMCFFJFCDYUDEZWEFCZDPJFJWOLDEFOEDOZWDDYZEFJLLDEKCLYDWEUDO-
DZCKCFJDOZWDJTTJOKEOCFCUELYDJUDZDCPVCSSCFCUEELODYCLJUCKDMFJQCYYJOSJUJOEO-
CXWCPQEDYLDEFOEZDPJFJYEFJKCDYWVDCKCLYDPZWKDMFJOQWUJQDXWJOKEYCFJKCLYDWEUDODJF-
FDGCFJJYZFEYYEUJOEOCXWCPQEDYLDEFOEXWCPQELLDZEUTJQQDJUEVCFQWQQEZDEZBCFDQCOC-
QCZJFEPWXWCPQJTCYYJQCFFJGDOGDQEJFCPDPQCFCWEUDODKCYYEGCPQ
```

Sul blog Research in Action è possibile utilizzare la libreria sviluppata con Blockly nel corso delle attività del progetto: http://researchinaction.it/myblockly/criptography.html con cui provare a usare i metodi di cifratura e di analisi affrontati.
In particolare le funzioni *ContaOccorrenzeCarattere* e *ContaOccorrenzeAlfabeto* permettono di contare il numero di occorrenze di un carattere specifico o di un insieme di caratteri in un dato testo.
Blockly è un generatore di codice open source sostenuto da Google: https://developers.google.com/blockly/.

Il tuo compito è quello di decifrare questo testo. A questo punto può essere utile qualche strumento in più **(cfr. Un suggerimento a pagina 20)**.

Per quanto possibile, cercheremo di usare il termine simboli, come abbiamo fatto finora, per le lettere dell'alfabeto cifrato (che scriveremo in maiuscolo, come di consueto) riservando la parola lettere per l'alfabeto in chiaro (che sarà riportato in minuscolo).

Per prima cosa determina la frequenza con cui ogni simbolo appare nel testo cifrato.

Nota a margine: il sito *Parole con* https://www.parolecon.it/ può essere di aiuto: permette infatti di cercare le parole italiane che iniziano, finiscono o contengono una determinata stringa di caratteri.

Completa la tabella che segue (che in seguito chiameremo *Tabella di analisi*) inserendo in corrispondenza di ogni carattere la sua frequenza in percentuale (nella terza e settima riga, quelle che iniziano con il simbolo di percentuale).

Per comodità, nelle righe che iniziano con il simbolo # puoi inserire il numero di occorenze di un dato carattere, nelle righe che iniziano con % la percentuale sul totale. Per ora lascia in bianco la quarta e ottava riga della tabella, ci serviranno in seguito.

Tabella di analisi

	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
#													
%													

Adesso siamo pronti per iniziare il nostro lavoro da agente segreto!

Un avvertimento: come avrai notato le percentuali della tabella delle frequenze e quelle che hai ricavato (e digitato nella tabella di analisi) non sono esattamente uguali, il testo che stai analizzando è troppo corto e potrebbe avere una distribuzione diversa dalla media dei testi normalmente scritti in italiano. Per cui dovremo considerare questi numeri con molta attenzione. Per esempio, non è detto che un simbolo che nel nostro testo ha una frequenza di 0.73% in italiano sia meno frequente di un altro che ha una frequenza di 0.98%. Semplicemente, con ogni probabilità, sono le due lettere davvero poco usate ma niente di più.

A una prima occhiata dovresti notare cinque simboli che hanno una frequenza pari a zero (e quindi non sono presenti nel testo cifrato). Devono per forza corrispondere alle lettere che hanno una frequenza molto molto bassa in italiano. Cancella, sbarra in qualche modo le caselle corrispondenti della tabella di analisi.

Infatti, per decifrare il testo non siamo interessati a quali lettere corrispondono i simboli eliminati: non essendo presenti nel testo sono inutili dal nostro punto di vista.

Ora analizza ancora le frequenze della tabella di analisi.

Ci sono due simboli che hanno percentuali molto basse e che devono corrispondere alle lettere h e z che sono quelle che in italiano, tra le lettere rimast, hanno le frequenze minori. Prova a sostituire i due simboli in questione con le lettere *h* e *z* nel testo cifrato.

Bada bene, per quanto detto sulle frequenze calcolate, devi provare entrambe le combinazioni e cercare di escludere una delle due analizzando il risultato dopo la sostituzione. È molto probabile che, per qualche motivo, una delle due ipotesi non sia accettabile. Per esempio, alcune consonanti, in italiano, non sono mai doppie.

Nella tabella di analisi puoi usare le due righe vuote per inserire, via via, le lettere corrispondenti a ciscun simbolo.

Se a questo punto sei in difficoltà niente paura, il compito non è facile. Prova a leggere la soluzione di questo quesito prima di procedere oltre **(cfr. 7.2 La criptoanalisi a pagina 22)**.

Adesso concentriamo l'attenzione sui simboli che nel testo cifrato hanno la frequenza più alta. Pare chiaro che devono corrispondere alle vocali *a*, *e*, *i* che sono le lettere più usate in italiano.

Prova a sostituire le lettere *e*, *i* ai simboli più frequenti. Anche in questo caso devi esplorare entrambe le possibilità ed escluderne una.

Tieni conto che alcune combinazioni di lettere, di vocali in questo caso, in italiano non sono possibili.

Leggi con cura il testo dopo la sostituzione (la migliore delle due che hai provato), c'è una sequenza di lettere (ancora solo vocali) che suggerisce una parola, una delle poche che contiene quella specifica sequenza?

Se è così puoi interpretare gli altri simboli dell'ipotetica parola e, magari, riuscire a decifrare una consonante.

Sostituisci la consonante e inizia a separare (aggiungere spazi) le parole che riconosci.

In particolare, controlla se ci sono vocali doppie: vuol dire che si tratta di lettere che quasi sicuramente appartengono a parole diverse (lettera finale per una parola, iniziale per l'altra), puoi inserire anche qui spazi per separare le due parole.

Ora di nuovo. Considera i due simboli con frequenza più alta tra quelli rimasti, devono corrispondere alle vocali *a*, *o* che in italiano sono molto usate. Prova a sostituire i simboli individuati con queste due lettere.

Ancora una volta hai davanti due possibili strade, cerca di escludere una delle due. Ricorda che alcune sequenze di lettere (ancora vocali, in questo caso) in italiano non sono possibili o sono molto molto rare.

Una volta sostituite le due vocali ai simboli corrispondenti, leggi il testo e cerca di riconoscere parole o embrioni di parole, in particolare controlla che non ci siano doppie vocali. Separa queste parole con degli spazi.

Se hai già riconosciuto alcune parole, le sillabe che le precedono potrebbero essere articoli.

Leggi ancora il testo, riconosci possibili articoli? Se è così dovresti riuscire a decifrare almeno un simbolo che corrisponde a una vocale.

Avrai notato che alcuni simboli appaiono in coppia. Con ogni probabilità sono consonanti.

Usando le frequenze, cerca le possibili consonanti (che in italiano possono apparire anche come doppie) che hanno frequenze più alte e paragonabili a simboli che nel testo cifrato appaiono anch'essi come doppi.

A questo punto non possiamo più aiutarti. Le possibili strade per decifrare il testo crescono in modo esponenziale. Pian piano dovresti iniziare a riconoscere alcune parole o *quasi-parole* che possono essere usate per interpretare un altro simbolo, il nuovo simbolo permetterà di riconoscere un'altra parola, e così via ...

In ogni caso, le soluzioni riportano un possibile percorso di analisi del testo cifrato, fino alla sua decrittazione completa **(il testo in chiaro lo trovi a pagina 24)**.

Anche in questo caso concludiamo riflettendo su un aspetto meno operativo ma ugualmente importante.

Cosa è necessario comunicare al destinatario, oltre al messaggio stesso naturalmente, perché riesca a decifrare correttamente e senza sforzo il testo cifrato.

4. Cifratura con frase chiave

Il metodo di cifratura con frase chiave è ancora un algoritmo che opera sostituendo a ogni lettera un determinato simbolo, in modo simile a quanto appena visto, ma che cerca di rendere più facile la trasmissione della chiave al destinatario.

In pratica la chiave è una frase, quasi sempre di senso compiuto (e quindi facile da ricordare o da intepretare), che viene usata per costruire l'alfabeto cifrato. Per esempio, se la nostra chiave è *le leggi cambiano ma le cose giuste restano uguali*, cerchiamo di costruire una tabella per la sostituzione molto simile a quelle usate negli esempi precedenti.

Come di consueto, sulla prima riga mettiamo le lettere dell'alfabeto in chiaro (e quindi in minuscolo) ordinate. La seconda riga conterrà le corrispondenti lettere cifrate (e quindi in maiuscolo).

Digitiamo ora le lettere della frase chiave una per casella partendo dalla prima, quella corrispondente alla lettera *a*. Evitiamo di ripetere le lettere già inserite per cui, per esempio, la *L* di *LEGGI* non appare perché la stessa lettera è già stata inserita in prima posizione (quando abbiamo inserito l'articolo *LE*) così come non appare la *E* sempre di *LEGGI*. Le lettere appena inserite sono evidenziate in colore rosso.

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
L	E	G	I	C	A	M	B	N	O	S	U	T	R												

Ora completiamo la tabella inserendo le restanti lettere dell'alfabeto ancora mancanti subito dopo l'ultima lettera della frase chiave proseguendo in ordine alfabetico (senza ripetere, ovviamente, simboli già presenti nella tabella) eventualmente iniziando di nuovo dall'inizio dopo aver aggiunto la zeta.

Sul blog Research in Action è possibile utilizzare la libreria sviluppata con Blockly nel corso delle attività del progetto: http://researchinaction.it/myblockly/criptography.html con cui provare a usare i metodi di cifratura e di analisi affrontati. In particolare le funzioni *FraseChiaveCifra* e *FraseChiaveDecifra* permettono di cifrare e decifrare un testo fornendo una frase come chiave.
Blockly è un generatore di codice open source sostenuto da Google: https://developers.google.com/blockly/.

In altre parole, l'ultima lettera inserita della frase chiave è la *R* quindi proseguiamo con la *S*, non inseriamo la *T* e la *U* perché già presenti (al dodicesimo e al tredicesimo posto), e così via. Arrivati alla *Z* ricominciamo da capo: non inseriamo *A*, *B* e *C* (già presenti), la *D* invece sì, saltiamo la *E* (è già al secondo posto) e proseguiamo con *F*, *H*, *J*, K e infine *P*. A questo punto abbiamo completato la seconda riga della tabella e ogni lettera in chiaro ha una corrispondente cifrata. Anche in questo caso le lettere inserite per ultime sono evidenziate in colore rosso.

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

La tabella è completa e possiamo passare a cifrare il testo. Operazione che si esegue esattamente nello stesso modo che abbiamo già visto.

4.1. Proviamo

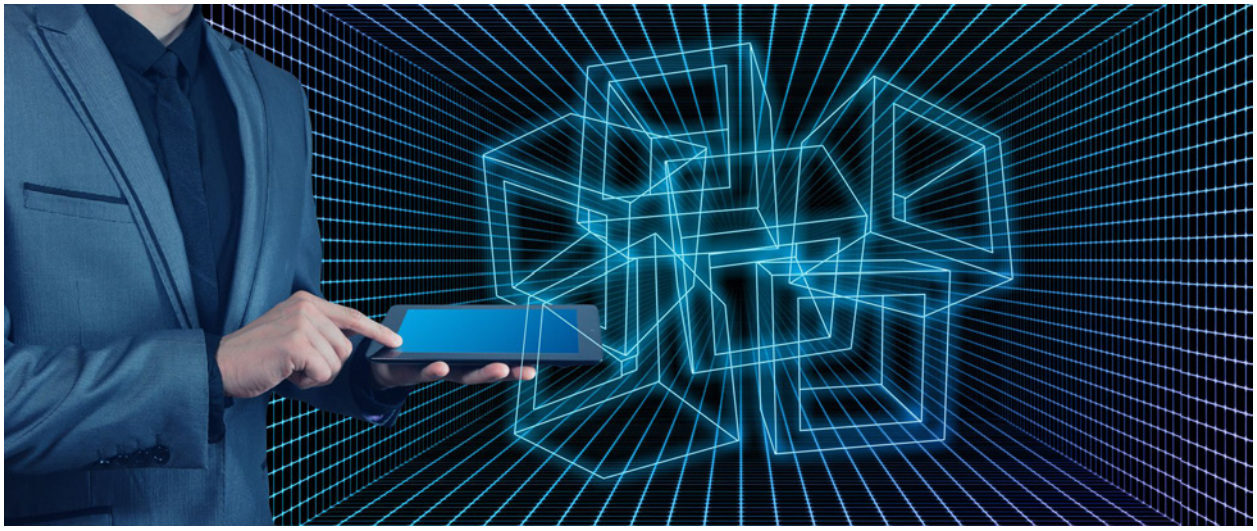
Scegli una frase, non è necessario che sia particolarmente originale, da usare come chiave e crea un alfabeto cifrato nel modo che abbiamo appena descritto.

In altre parole, costruisci la (ormai solita) tabella. Ti sarà utile per la cifratura e decifratura dei messaggi. Puoi usare la tabella che trovi qui di seguito.

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

Prova a cifrare e decifrare la frase *Due cose sono infinite: l'universo e la stupidità umana, ma riguardo l'universo ho ancora dei dubbi* **con il metodo che abbiamo appena appreso.**

Ricorda sempre di eliminare tutti i caratteri che non sono alfanumerici prima della cifratura.



4.2. Criptoanalisi

Come in precedenza, mettiamoci nei panni di una spia e proviamo a decifrare il messaggio senza conoscere la chiave, le cose non sono così semplici anzi, diventano veramente complesse come abbiamo visto.

Abbiamo cifrato la frase che segue con il metodo della frase chiave, prova a decifrare il testo e a ricostruire l'alfabeto cifrato.

QRIFVFVCIBEVRUWURQPIGICCUFEIRQJUGVWUCIPUJJINVKPINIJIKEUKPILECEPIHQEFVJPIR-
GVOEVWURPQSVWESURPVKFEJEPVGIWWURPQJIKERNVRPJIRVFUJCUKPJIGUQVSEREUGVRRULURR-
QPJEPEJIOIBBEUJIOIBBUNDUKIRRVJEGUJUUDIRRVJEFJUKVECOQKPVFUJCUAUKPUNISFUKPJEKU-
KENVRPICIWUNNDEIFVFVCIBEVRUEJJENVRVKNELECUGIHQIRGVWEWURUCCIJSVREIUERQVWEWUR-
QPEFEQGEGEUNESECIFUJKVRUGUWVRVCICVJVAUCENEPIIUCUBUIJGLVQAAEUJ

Qualche suggerimento:

» elimina subito dalla tabella di analisi i simboli che hanno frequenze nulle o quasi (corrispondono alle lettere dell'alfabeto anglosassone che non sono - quasi - mai usate in italiano);
» confronta le frequenze più basse dei simboli che rimangono con le frequenze dell'alfabeto in chiaro, cerca, magari con una o più prove, di individuare alcune corrispondenze;
» passa ora alle frequenze più alte, se devi scegliere controlla che non ci siano doppie vocali;
» le doppie vocali in italiano sono rarissime mentre le doppie consonanti molto frequenti, quindi con ogni probabilità coppie di simboli successivi nel testo cifrato saranno consonanti, usa le frequenze per cercare di capire quali;
» leggi ogni volta il testo con attenzione cercando di individuare o riconoscere parole o espressioni;
» fai ogni volta una copia del testo cifrato e conserva i tentativi, capita spesso di dover tornare indietro e rivedere scelte già fatte.

Per questo lavoro puoi usare la tabella che trovi qui di seguito.

Tabella di analisi

	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
%													

5. Metodo di Vigenère

Come abbiamo visto, non è affatto impossibile decifrare un testo criptato con uno dei metodi di sostituzione che abbiamo affrontato. In particolare, la debolezza sta nella frequenza dei simboli che, per forza di cosa, rispecchia la frequenza delle lettere dell'alfabeto in chiaro, infatti una certa lettera in un messaggio è cifrata sempre con lo stesso simbolo e, viceversa, un simbolo rappresenta sempre la stessa lettera.

Molto più difficile sarebbe se nello stesso messaggio la stessa lettera fosse rappresentata da simboli diversi: una delle armi migliori dei crittoanalisti (l'analisi delle frequenze) d'un tratto verrebbe meno. Questa è l'idea su cui è basato il *metodo di Vigenére.*

Il trattato di Blaise de Vigenère (*Traicté de Chiffres*, pubblicato nel 1586) è basato su un'idea di Leon Battista Alberti, successivamente sviluppata da Joannes Trithemius negli anni successivi e da Gianbattista della Porta nel secolo successivo. L'idea base è quella di usare il metodo di sostituzione cambiando via via l'alfabeto cifrato in modo da impedire l'analisi delle frequenze.

Si tratta di cifrare un messaggio usando 26 alfabeti cifrati diversi usando il meccanismo di Cesare che abbiamo già visto ma scegliendo via via tra i 26 alfabeti cifrati. Per cifrare o decifrare un testo si costruisce innanzitutto la *tavola di Vigenère* riportando nella prima riga l'alfabeto in chiaro e nelle 26 righe successive altrettanti alfabeti cifrati costruiti con il metodo di Cesare e con chiave aumentata di un'unità per ogni riga (a partire da uno).

Tabella di Vigenère

	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
1	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
2	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	B
3	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
4	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	D
5	F	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
6	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
7	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	G
8	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
9	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	I
10	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
11	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
12	M	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
13	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
14	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
15	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	O
16	Q	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
17	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
18	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
19	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	S
20	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
21	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
22	W	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
23	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
24	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	X
25	Z	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
26	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

In questa tabella la prima colonna riposrta la chiave usata per costruire l'alfabeto cifrato che si trova nella stessa riga.

A questo punto si costruisce la tabella per la cifratura: nella seconda riga viene riportato il messaggio da cifrare mentre sulla prima riga una frase chiave (scelta in modo opportuno, abbastanza lunga rispetto al messaggio e che contenga molte lettere differenti), eventualmente ripetuta, in modo che ogni lettera della frase chiave (e delle sue ripetizioni) corrisponda a una lettera del

messaggio. Per esempio, volendo cifrare il testo *La mente è come un paracadute* con la frase chiave *Funziona solo se si apre*, la tabella sarebbe:

f	u	n	z	i	o	n	a	s	o	l	o	s	e	s	i	a	p	r	e	f	u	n	z
l	a	m	e	n	t	e	e	c	o	m	e	u	n	p	a	r	a	c	a	d	u	t	e

Notare che la frase chiave è in parte ripetuta in modo da avere nella prima riga tanti caratteri quanti ce ne sono nella seconda (che contiene il messaggio da cifrare).

Bene! Ora la lettera che si trova nella prima riga indica quale alfabeto usare per cifrare la corrispondente lettera del messaggio (quello che si trova nella riga che inizia proprio con questa lettera). In pratica:

» La prima lettera del messaggio è *l*, la lettera corrispondente della frase chiave è *f*, vuol dire che dobbiamo usare la riga numero 5 della tabella di Vigenère (quella che contiene l'alfabeto che inizia con *F*) e quindi la lettera cifrata è *Q* (il simbolo che corrisponde alla *l* nell'alfabeto cifrato usato).

» La seconda lettera da cifrare *a*, la corrispondente nella chiave è *u*, quindi l'alfabeto si trova nella ventesima riga (quella che inizia per *U*) e il simbolo cifrato è proprio *U* (il simbolo corrispondente alla lettera a in questo alfabeto).

Qui sopra la scelta del primo simbolo, quello usato per cifrare la lettera elle. Nella tabella che si trova nella pagina precedente le stesse due righe usate sono evidenziate in colore verde e i simboli sostituiti in colore rosso.

Proseguendo in questo modo abbiamo *QUZDVHREUCXSMRHIRPTEIOGD*.

È importante notare che, per esempio, le due lettere *e* della parola *mente* sono cifrate con due simboli differenti: *D* e *R*. Viceversa, i due simboli *H* (presenti al sesto e al quindicesimo posto del testo cifrato) sostituiscono due lettere diverse, rispettivamente la *t* e la *p*: addio all'analisi delle frequenze!

Per decifrazione procediamo in modo simile, costruiamo una tabella con la frase chiave (eventualmente ripetuta) nella prima riga e la frase cifrata nella seconda riga.

f	u	n	z	i	o	n	a	s	o	l	o	s	e	s	i	a	p	r	e	f	u	n	z
Q	U	Z	D	V	H	R	E	U	C	X	S	M	R	H	I	R	P	T	E	I	O	G	D

A questo punto, per ciascun simbolo da decifrare, cerchiamo nella tabella di Vigenére la riga indicata dalla lettera della frase chiave corrispondente e la colonna, in questa riga, contenente il simbolo da decifrare: la lettera che si trova nella prima riga della tabella e nella stessa colonna è quello che ci serve. Per esempio:

» Iniziamo dal simbolo *Q*: la lettera corrispondente, nella frase chiave, è la *f* che ci porta alla quinta riga della tabella. In corrispondenza della *Q*, nella prima riga troviamo la lettera *l*. abbiamo decifrato il primo simbolo.

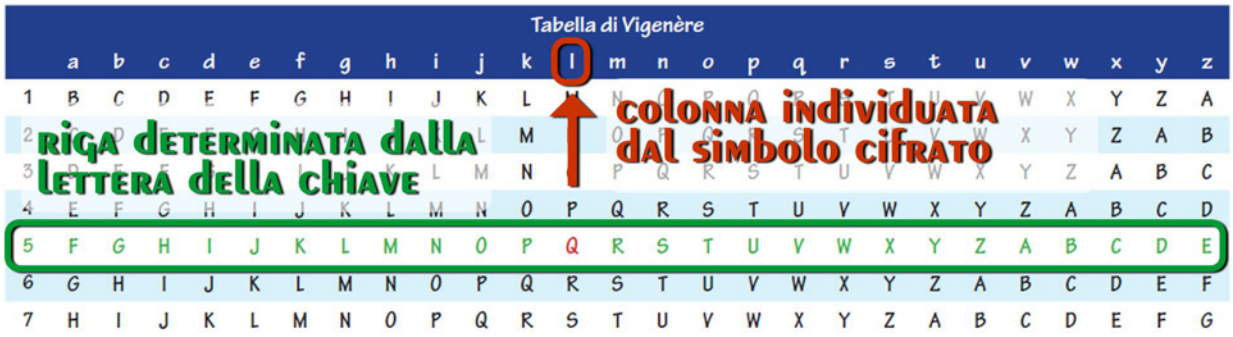
» Passiamo al secondo, la *U*. La lettera corrispondente è la *u* che ci indica la ventesima riga della

Blockly Google
A Visual Programming Language

Sul blog Research in Action è possibile utilizzare la libreria sviluppata con Blockly nel corso delle attività del progetto: http://researchinaction.it/myblockly/criptography.html con cui provare a usare i metodi di cifratura e di analisi affrontati. In particolare le funzioni *VigenereCifra* e *VigenereDecifra* permettono di cifrare e decifrare un testo fornendo una frase come chiave.
Blockly è un generatore di codice open source sostenuto da Google: https://developers.google.com/blockly/.

Curiosamente, l'anno di pubblicazione del trattato di Vigenère è lo stesso del processo a Mary Stuart. La regina di Scozia fu condannata alla pena capitale anche grazie alle decifrazione della sua corrispondenza segreta con Walsingham, cifrata con un metodo troppo facile da decifrare. Se lei o il suo segretario avessero usato la cifratura di Vigenère chissà se la storia avrebbe preso un altro corso.

tabella. Nella prima riga, in corrispondenza del simbolo da decifrare, troviamo la lettera a.



Proseguendo in questo modo otteniamo la frase originale *lameneteecomeunparacadute*: *la mente è come un paracadute.*

5.1. A questo punto tocca a te

Scegli una frase chiave e cifra, con il metodo di Vigenère, la frase *Il denaro non cura la malattia, solo i sintomi.* **Successivamente decifra la frase.**

Appare chiaro che la criptoanalisi di un testo cifrato con questo metodo diventa davvero difficile e non bastano più carta e penna ma sono necessari mezzi e tecnologie decisamente superiori. Per l'epoca era davvero un codice difficile fa forzare!

Malgrado sia stato usato in origine nel modo che abbiamo esposto, il metodo di Vigenère funziona anche senza costruire la tabella. Infatti è possibile determinare il simbolo da sostituire alla lettera del testo in chiaro solo attraverso semplici calcoli (le richieste che concludono questo capitolo sono abbastanza complesse, se le ritieni troppo difficili puoi passare direttamente al capitolo successivo - ... La macchina Enigma ... a pagina ...) oppure leggere le soluzioni che abbiamo proposto (cfr. ... Il meto di Vigenère a pagina ...).

Prova a determinare un algoritmo, una procedura, per cifrare un testo usando il metodo di Vigenère senza far riferimento alla tabella omonima.

Un suggerimento: prova a confrontare il numero della riga e quello della colonna del carattere cifrato con la posizione dello stesso carattere nell'alfabeto in uso.

Cerchiamo di spiegare meglio, nell'esempio che abbiamo fatto la *c* di *come* è cifrata con il simbolo *U* (vedi anche la figura qui sopra che riporta parte della tabella di Vigenère).

Se ripetiamo il procedimento di cifratura ci accorgiamo che il simbolo si trovava nella *18° riga* (quella corrispondente al carattere *s* della frase chiave) e nella *terza colonna*: questi sono i numeri

di riga e colonna a cui facevamo riferimento prima. La lettera *U* nell'alfabeto anglosassone si trova al 21° posto: questa è la posizione nell'alfabeto che abbiamo citato in precedenza.

Il Toolbox: http://researchinaction.it/wp-content/uploads/2018/11/00-Toolbox.pdf introduce i principali concetti e metodi per un approccio informale al coding e alla formalizzazione delle procedure e degli algoritmi (cfr. 6. Algoritmi a pagina 23).

C'è un legame tra 18 e 3 e 21? Questa relazione vale per ogni altro simbolo del testo cifrato?

Una volta formalizzato un algoritmo per la cifratura (compito niente affatto facile), dovrebbe essere più semplice proseguire sulla stessa strada.

Determina un algoritmo, una procedura, per decifrare un testo che è stato cifrato con il metodo di Vigenère e come in precedenza, fai in modo che l'algoritmo non abbia bisogno della tabella.

Deve valere una relazione simile a quella determinata in precedenza, che abbiamo cercato di introdurre nel suggerimento. Per formalizzare l'algoritmo puoi usare il *foglio di carta* che trovi in fondo a questa pagina.

Un suggerimento

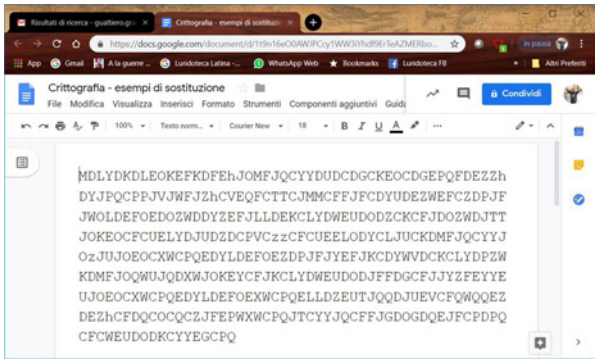
Le operazioni di cifratura, in modo particolare quando il testo è particolarmente lungo, sono piuttosto tediose e fonte di ripetuti errori se eseguite manualmente.

Per questo laboratorio consigliamo di usare un'applicazione per gestire testi (come per esempio *Microsoft Word* o *Google Documenti*) perchè tutte queste applicazioni hanno funzioni di sostituzione che possono rendere semi-automatiche le operazioni di cifratura.

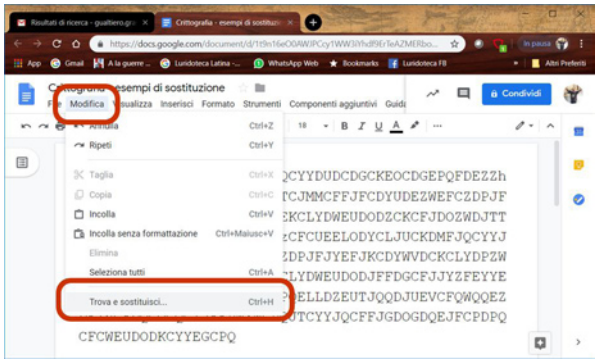
Qui di seguito un esempio di come è possibile decifrare (ma per la cifratura la procedura è la stessa) un testo usando *Google Documenti*.

Useremo, come in tutto il resto del laboratorio, le solite convenzioni: testo cifrato in maiuscolo e testo in chiaro in minuscolo.

Nella figura che segue si può vedere un documento *Google Documenti* in cui abbiamo importato (mediante le funzionalità di *copia-incolla*) uno dei testi cifrati che abbiamo usato in precedenza **(cfr. 7.2 La criptoanalisi a pagina 24)**.



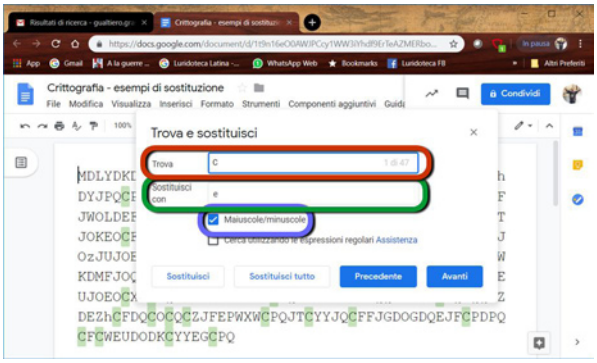
Vogliamo sostituire la lettera *e* (dell'alfabeto in chiaro) al simbolo *C* (dell'alfabeto cifrato). Selezionamo la voce *Trova e sostituisci* del menu *Modifica*, come mostrato nella figura che segue (il menu e la voce che ci serve sono evidenziati dai *box* di colore rosso):



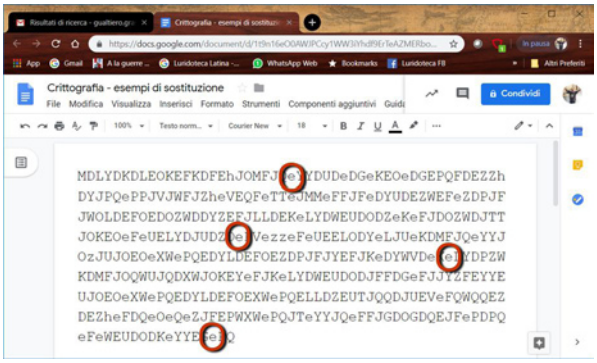
Selezionando la voce del menu si apre una finestra di dialogo in cui specificare il testo da sostituire e scegliere alcune altre opzioni delle quali una per noi è particolarmente utile. Come si vede dalla figura che segue abbiamo:

» digitato la lettera *C* (attenzione: è importante che sia digitata in maiuscolo) nel campo *Trova* (evidenziato dal *box* di colore rosso): si tratta del testo che verrà sostituito;
» digitato la lettera *e* (attenzione: in minuscolo) nella casella *Sostituisci con* (evidenziata dal *box* di colore verde): si tratta del testo che sostituirà il precedente;
» selezionato l'opzione *Maiuscole/minuscole* (evidenziata dal box di colore azzurro.

Si intravede nel testo la selezione di tutte le lettere *C* maiuscole: sono le lettere su cui opererà la sostituzione.



Se a questo punto premiamo il pulsante *Sostituisci tutto*, chiediamo all'applicazione di sostituire a ogni lettera *C* maiuscola una *e* minuscola come si vede dalla figura che segue (alcune delle *e* sono evidenziate in colore rosso). Il pulsante *Sostituisci*, invece, sostituisce una lettera alla volta, permettendo all'utente di scegliere via via se procedere oppure no.



La selezione dell'opzione *Maiuscole/minuscole* è importante perchè in questo modo si evita di sostituire testo già in chiaro (in minuscolo)!

Altre applicazioni permettono la sostituzione solo nel testo selezionato: questo rende più facile la gestione di versioni successive del messaggio nello stesso documento, via via che si procede nell'analisi del testo.

Soluzioni

Crittografia 1 - L'arte di svelare segreti!

6. Il metodo di Cesare

6.1. Proviamo

Scegli una chiave (in sostanza un numero intero) e prova a cifrare e decifrare la frase che segue: *Due cose sono infinite: l'universo e la stupidità umana, ma riguardo l'universo ho ancora dei dubbi.* **Ricorda di eliminare tutti i caratteri che non sono alfanumerici prima della cifratura.**

Sul blog Research in Action è possibile utilizzare la libreria sviluppata con Blockly nel corso delle attività del progetto: http://researchinaction.it/myblockly/criptography.html con cui provare a usare i metodi di cifratura e di analisi affrontati. In particolare le funzioni *CodiceCesareCifra* e *CodiceCesareDecifra* permettono di cifrare e decifrare un testo con il metodo di Cesare specificando la chiave e il testo in chiaro.
Blockly è un generatore di codice open source sostenuto da Google: https://developers.google.com/blockly/.

Prima di tutto ripuliamo il testo dai caratteri non alfanumerici, otteniamo la stringa: *duecosesonoinfiniteluniversoelastupiditaumanamariguardoluniversohoancoradeidubbi.*

Scegliamo come chiave $c = 8$. Questo vuol dire che a ogni lettera dell'alfabeto in chiaro sarà sostituita quella che si trova otto posizioni *dopo*. La tabella per la cifratura quindi sarà simile a questa che segue dove abbiamo inserito la lettera *I* al primo posto perchè si tratta della lettera che si trova otto posizioni dopo la a nell'alfabeto anglosassone, e poi abbiamo proseguito scrivendo tutti gli altri simboli, uno per casella, in ordine alfabetico. Arrivati alla *Z* abbiamo ri-iniziato dalla *A*.

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
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

La frase cifrata è quindi *LCMKWAMAWVWQVNQVQBMTCVQDMZAWMTIABCXQLQBICUIVIUIZQOCIZLWTCVQDMZAWPWIVKWZILMQLCJJQ.*

Sei riuscito a decifrare il messaggio segreto? Come valuti la sicurezza del codice di Cesare?

Se hai provato a decifrare un messaggio *segreto* creato da qualcun altro (e quindi senza conoscerne la chiave), come richiesto dal laboratorio ti sarai accorto che ci sono solo 26 possibili chiavi anzi, in realtà 25 perchè la chiave $c = 26$ equivale a non cambiare nulla, a ogni lettera dell'alfabeto in chiaro corrisponde la stessa lettera.

Un messaggio cifrato con questo metodo, quindi, richiede al più 25 tentativi per capire qual è la chiave. Per esempio, il testo *EJCQANNWCHEARAJPEEILKNPWJPEOKJKEHNEOQHPWPKZEYWQOAXWJWHE* è stato cifrato con il metodo di cesare, se proviamo via via tutte le chiavi, magari partendo da 25 invece che da uno otteniamo:

chiave	testo decifrato
25	fkdrbooxdifbsbkqffjmloqxkqfplklfiofpriqxqlafzxrpbyxkxif
24	glescppyejgctclrggknmprylrgqmlmqjpgqsjryrmbgaysqczylyjg
23	hmftdqqzfkhdudmshhlonqszmshrnmnhkqhrtkszsnchbztrdazmzkh
22	inguerraglieventiimportantisonoilrisultatodicausebanali

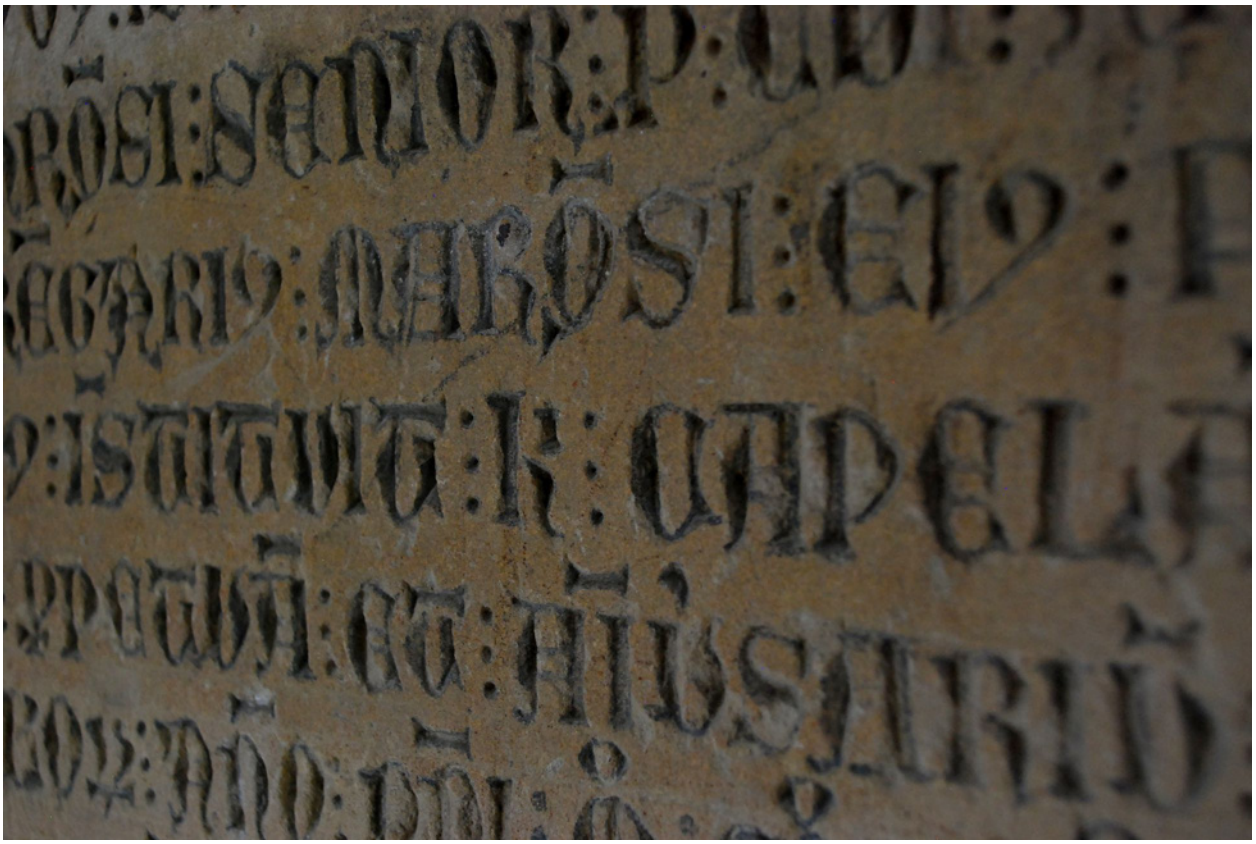
E quindi appare chiaro che la chiave deve essere $c = 22$ e il testo in chiaro, aggiungendo di nuovo gli spazi e le lettere accentate è: *in guerra gli eventi importanti sono il risultato di cause banali*! Tra l'altro proprio una frase attribuita a Giulio Cesare.

Anche tentando a partire dalla chiave $c = 1$ saremmo arrivati allo stesso risultato ma con un nu-

mero maggiore di tentativi. In realtà, lavorando in gruppo e dividendosi le chiavi *un po' per uno* si può decifrare un testo senza troppa fatica. Il problema è che le possibili sostituzioni son troppo poche!

Ancora meno tempo se si usa un dispositivo elettronico programmato per tentare, una dopo l'altra le chiavi da 1 a 25 (vedi la nota a margine qui accanto). Un giudizio sulla sicurezza del metodo di Cesare non può che essere negativa, a oggi è semplicemente un passatempo, non certo un procedimento con cui garantire la sicurezza dei messaggi. **Cosa è necessario comunicare al destinatario, oltre al messaggio stesso naturalmente, perché riesca a decifrare correttamente e senza sforzo il testo cifrato.**

Malgrado tutto è stato utile affrontare questo metodo. Appare chiaro infatti che, in qualche modo, il destinatario deve conoscere, oltre al messaggio ciffrato, anche la chiave usata per ciffrarlo. A meno di non voler tentare tutte le possibili scelte!



7. Cifratura per sostituzione

7.1. Proviamo

Scegli una chiave. In altre parole, costruisci una tabella disponendo casualmente i simboli della seconda riga.

Prova a cifrare e decifrare la frase *Due cose sono infinite: l'universo e la stupidità umana, ma riguardo l'universo ho ancora dei dubbi* **con il metodo che abbiamo appena appreso. Ricorda sempre di eliminare tutti i caratteri che non sono alfanumerici prima della cifratura.**

Con ogni probabilità la chiave che abbiamo scelto qui sarà differente da quella che hai usato ma non è così importante, importante è comprendere il funzionamento di questo algoritmo. Allora, scegliamo come chiave *IGBSMFALTKZQEDJRPWYVONXCUH*. Costruiamo la tabella che ci aiuterà nella cifratura inserendo nella prima riga l'alfabeto in chiaro e nella seconda la nostra chiave.

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

Il testo da cifrare, ripulito dei caratteri non necessari è *duecosesonoinfiniteluniversoelastupiditaumanamariguardoluniversohoancoradeidubbi*. Sostituendo alle lettere del testo in chiaro i corrispondenti simboli si ottiene *SOMBJYMYJDJTDFTDTVMQODTNMWYJMQIYVORTSTVIOEIDIEIWTAOIWSJQODTNMWYJLJIDBJWISMTSOGGT*.

Per decifrare questo testo possiamo usare ancora la tabellina che abbiamo preparato. Per esempio, il primo simbolo del testo cifrato è *S*: lo cerchiamo nella seconda riga della tabella e vediamo che corrisponde alla lettera *d*. Il secondo simbolo è *O* che corrisponde alla lettera *u*. E così via ... (abbiamo evidenziato in colore rosso i due simboli sostituiti nella tabella precedente). Il messaggio, una volta decifrato, è *duecosesonoinfiniteluniversoelastupiditaumanamariguardoluniversohoancoradeidubbi* che, aggiungendo di nuovo gli spazi e le lettere accentate diventa proprio *due cose sono infinite: l'universo e la stupidità umana, ma riguardo l'universo ho ancora dei dubbi.*

Chiedi a qualcuno di cifrare un messaggio con una chiave a sua scelta, realizzata come appena descritto, e cerca di decifrarlo senza conoscerne la chiave. Quante possibili chiavi ha potuto scegliere il creatore del messaggio cifrato?

Iniziamo con rispondere all'ultima parte del quesito: le possibili chiavi sono veramente tante: *403.291.461.126.605.635.584.000.000*, oltre quattrocento milioni di miliardi di miliardi. Corrispondono infatti alle permutazioni delle 26 lettere dell'alfabeto anglosassone. È davvero impossibile forzare il codive usando la *forza bruta* (provando le chiavi una a una come abbiamo fatto per il metodo di Cesare).

Sei riuscito? Come valuti la sicurezza di questo algoritmo di cifratura rispetto al precedente? Può garantire la sicurezza?

A questo punto verrebbe da dire che si tratta di un metodo abbastanza sicuro. A meno di non avere un computer superpotente, in grado di provare una chiave in un miliardesimo di miliardesimo di secondo, sembra impossibile riuscire a decifrare il messaggio. Ma ... forse dove non arriva la forza, può arrivare l'intelligenza.

Cosa è necessario comunicare al destinatario, oltre al messaggio stesso naturalmente, perché riesca a decifrare correttamente e senza sforzo il testo cifrato.

Blockly Google
A Visual Programming Language

Sul blog Research in Action è possibile utilizzare la libreria sviluppata con Blockly nel corso delle attività del progetto: http://researchinaction.it/myblockly/criptography.html con cui provare a usare i metodi di cifratura e di analisi affrontati. In particolare la funzione *CreaChiaveCasuale* genera una chiave per il metodo di sostituzione del tutto casuale e le funzioni *SostituzioneCifra* e *SostituzioneDecifra* permettono di cifrare e decifrare un testo con il metodo di sostituzione specificando la chiave e il testo in chiaro.
Blockly è un generatore di codice open source sostenuto da Google: https://developers.google.com/blockly/.

Come in precedenza anzi, più di prima, è assolutamente necessario che il destinatario conosca la chiave. E qui il problema inizia a farsi complicato, sempre di più la trasmissione sicura di un messaggio richiede il trasferimento del messaggio ma anche della chiave!

7.2. La criptoanalisi

Proviamo a forzare il codice con cui è stato cifrato il testo.

Completa la tabella che segue (che in seguito chiameremo *Tabella di analisi*) inserendo in corrispondenza di ogni carattere la sua frequenza in percentuale (nella terza e settima riga, quelle che iniziano con il simbolo di percentuale).

Contando le occorrenze di ogni simbolo nel testo otteniamo quanto segue (le righe con il simbolo # contengono il numero di occorrenze del simbolo corrispondente, le righe con il simbolo % la percentuale di queste occorrenze: la frequenza con cui il simbolo appare nel testo):

Tabella di analisi

	A	B	C	D	E	F	G	H	I	J	K	L	M
#	0	4	47	51	40	34	6	0	0	38	14	15	6
%	0,00	0,98	11,46	12,44	9,76	8,29	1,46	0,00	0,00	9,27	3,41	3,66	1,46
		z	e	i									
	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
#	0	24	16	22	0	3	6	14	5	18	5	25	17
%	0,00	5,85	3,90	5,37	0,00	0,73	1,46	3,41	1,22	4,39	1,22	6,10	4,15
						h							

Nel seguito il simbolo -> indicherà la corrispondenza tra un determinato simbolo dell'alfabeto cifrato con la corrispondente lettera dell'alfabeto in chiaro. Per esempio, Q -> t indica che il simbolo *Q* corrisponde alla lettera *t* e che per cifrare il testo si è sostuita la *t* alla *Q*.

Inoltre, ogni volta che apporteremo modifiche al testo cifrato, cercheremo di evidenziare in colore rosso le porzioni interessanti per qualche motivo (perchè contengono stringhe interessanti, parole o *quasi-parole*, ...).

A una prima occhiata dovresti notare cinque simboli che hanno una frequenza pari a zero (e quindi non sono presenti nel testo cifrato). Devono per forza corrispondere alle lettere che hanno una frequenza molto molto bassa in italiano. Cancella, sbarra in qualche modo le caselle corrispondenti della tabella di analisi.

Nella tabella eliminiamo i simboli *A*, *H*, *I*, *N* e *R* che non sono presenti nel testo cifrato, devono corrispondere allle lettere *i*, *j*, *k*, *w* e *x* che in italiano sono rarissime (nella tabella abbiamo colorato in grigio le celle corrispondenti).

Ci sono due simboli che hanno percentuali molto basse e che devono corrispondere alle lettere h e z che sono quelle che in italiano, tra le lettere rimast, hanno le frequenze minori. Prova a sostituire i due simboli in questione con le lettere *h* e *z* nel testo cifrato.

I simboli *B* e *S* hanno ora le frequenze più basse, devono corrispondere alle lettere *h* e z. Potremmo avere *B* -> *z* e *S* -> *h* ma in questo caso la doppia S diventerebbe una doppia h e questo non è possibile. Quindi deve essere *B* -> *h* e *S* -> *z*, per cui il testo a questo punto è:

```
MDLYDKDLEOKEFKDFEhJOMFJQCYYDUDCDGCKEOCDGEPQFDEZZhDYJPQCPPJVJWFJZhCVE-
QFCTTCJMMCFFJFCDYUDEZWEFCZDPJFJWOLDEFOEDOZWDDYZEFJLLDEKCLYDWEUDO-
DZCKCFJDOZWDJTTJOKEOCFCUELYDJUDZDCPVCzzCFCUEELODYCLJUCKDMFJQCYYJOzJUJOEO-
CXWCPQEDYLDEFOEZDPJFJYEFJKCDYWVDCKCLYDPZWKDMFJOQWUJQDXWJOKEYCFJKCLYDWEUDODJF-
```

FDGCFJJYZFEYYEUJOEOCXWCPQEDYLDEFOEXWCPQELLDZEUTJQQDJUEVCFQWQQEZDEZhCFDQCOC-
QCZJFEPWXWCPQJTCYYJQCFFJGDOGDQEJFCPDPQCFCWEUDODKCYYEGCPQ

Nella tabella di analisi abbiamo inserito (ed evidenziato in rosso) le lettere *z* e *h* in corrispondenza dei simboli *B* e *S*. Le due z nel testo sono anch'esse evidenziate in colore rosso.

Prova a sostituire le lettere *e*, *i* ai simboli più frequenti. Anche in questo caso devi esplorare entrambe le possibilità ed escluderne una.

I simboli con la frequenza maggiore sono *C*, *D* e *J* che dovrebbero corrispondere alle vocali *a*, *e*, *i*. Dopo qualche tentativo ci sembra una buona idea *C* -> *e* e *D* -> *i* (nella tabella evidenziate in colore verde). Infatti la sostituzione contraria *C* -> *i* e *D* -> *e* produceva la sequenza *eie* che è davvero rara in italiano. Ora invece abbiamo la sequenza *iei* che potrebbe essere *miei* e quindi il simbolo *U* potrebbe essere la *m*. Ci accorgiamo anche che c'è una doppia *i* che possiamo separare aggiungendo uno spazio, così come possiamo separare la parola *miei*.

MiLYiKiLEOKEFKiFEhJOMFJQeYYi miei GeKEOeiGEPQFiEZZhiYJPQePPJVJWFJZheVEQFeTTeJM-
MeFFJFeiYmiEZWEFeZiPJFJWOLiEFOEiOZWi iYZEFJLLiEKeLYiWEmiOiZeKeFJiOZWiJTTJOKE-
OeFemELYiJmiZiePVezzeFemEELOiYeLJmeKiMFJQeYYJOzJmJOEOeXWePQEiYLiEFOEZiPJFJYE-
FJKeiYWVieKeLYiPZWKiMFJOQWmJQiXWJOKEYeFJKeLYiWEmiOiJFFiGeFJJYZFEYYEmJOEOeXWe-
PQEiYLiEFOEXWePQELLiZEmTJQQiJmEVeFQWQQEZiEZheFiQeOeQeZJFEPWXWePQJTeYYJQeFFJ-
GiOGiQEJFePiPQeFeWEmiOiKeYYEGePQ

Ora di nuovo. Considera i due simboli con frequenza più alta tra quelli rimasti, devono corrispondere alle vocali *a*, *o* che in italiano sono molto usate. Prova a sostituire i simboli individuati con queste due lettere.

I due simboli con frequenze maggiori rimasti sono *E* e *J* che devono corrispondere alle vocali *a* e *o*. Se sostituiamo *J* -> *o* viene fuori la combinazione iomi, rara, meglio allora sostituire *J* -> *a* e *E* -> *o*. Rileggendo il testo si trova la stringa *iYmio* dove *mio* potrebbe essere una parola e *iY* l'articolo e quindi *Y* -> *l*. Ora ci ritroviamo la stringa *olloma* che non può far parte della stessa parola, la separiamo in *ollo ma*.

MiLliKiLoOKoFKiFohaOMFaQelli miei GeKoOeiGoPQFioZZhilaPQePPaVaWFaZheVoQFeTTe-
aMMeFFaFe il mio ZWoFeZiPaFaWOLioFOoiOZWi il ZoFaLLioKeLliWomiOiZeKeFaiOZWiaT-
TaOKoOeFemoLliamiZiePVezzeFemo oLOileLameKiMFaQellaOzamaOoOeXWePQoilLioFOoZi-
PaFaloFaKeilWVieKeLliPZWKiMFaOQWmaQiXWaOKoleFaKeLliWomiOiaFFiGeFa alZFollo ma
OoOeXWePQoilLioFOoXWePQoLLiZomTaQQiamoVeFQWQQoZioZheFiQeOeQeZaFoPWXWePQaTel-
laQeFFaGiOGiQoaFePiPQeFeWomiOiKelloGePQ

Usando le frequenze, cerca le possibili consonanti (che in italiano possono apparire anche come doppie) che hanno frequenze più alte e paragonabili a simboli che nel testo cifrato appaiono anch'essi come doppi.

Il simbolo *F* deve essere una consonante e appare anche come doppia, un rapido confronto tra le frequenze delle consonanti in italiano e la tabella di analisi (dove la *F* ha ora la frequenza più alta) ci fa capire che corrisponde alla *r* o alla *n*, la sostituzione *F* -> *r* fa apparire *aMMerrare* che potrebbe essere *afferrare* e quindi *M* -> *f*. Ma anche *Qerra* potrebbe essere terra e fraQelli la parola *fratelli* per cui *Q* -> *t*. Rileggendo si trova *fratellaOSa* che potrebbe essere *fratellanza* e abbiamo altre due sostituzioni *O* -> *n* e *S* -> *z*. Ogni volta che individuiamo una parola separiamola aggiungendo gli spazi opportuni.

fiLliKiLonKorKirohan fratelli miei GeKo nei GoPtrioZZhilaPtePPaVaWraZheVotreTTe
afferrare il mio ZWoreZiParaWnLiornoinZWi il ZoraLLioKeLliWominiZeKerainZWiaT-
TanKoneremoLliamiZiePVezzeremo oLnileLameKi fratellanza manoneXWePtoilLiorno-
ZiParaloraKeilWVieKeLliPZWKifrantWmatiXWanKoleraKeLliWomini arriGera alZrollo
ma noneXWePtoilLiornoXWePtoLLiZomTattiamoVertWttoZioZheriteneteZaroPWXWePta-
Tella terra GinGitoarePiPtereWominiKelloGePt

Sta diventando sempre più facile individuare parole o *quasi-parole*. Se leggiamo con attenzione troviamo *VotreTTe* è quasi sicuramente *potrebbe*, *Womini* è *uomini*, *PVezzeremo*, visto che *V -> p*, è *spezzeremo* per cui *V -> p* come detto, *T -> b*, *W -> u*, *P -> s*.

A questo punto rileggiamo e separiamo le parole che si riconoscono facilmente come

fiLlidiLondordirohan fratelli miei Gedo nei Gostri occhi la stessa paura che potrebbe afferrare il mio cuore ci sarà un Liorno in cui il coraLLiodeLli uomini cederà in cui abbandoneremo Lliamicie spezzeremo oLnileLamedi fratellanza ma non è questo il Liorno ci sarà l'ora dei lupi e deLli scudi frantumati quando l'era deLli uomini arriGera al crollo ma non è questo il Liorno questoLLi combattiamo per tutto cio che ritenete caro su questa bella terra GinGitoa resistere uomini delloGest

Ormai il più è fatto. Dalle stringhe *Gedo* e *Gostri* si capisce che *G -> v*, da *Liorno* si intuisce anche che *L -> g*. Aggiungiamo qualche apostrofo dove necessario e abbiamo il nostro testo decifrato. Gran bel lavoro!

figli di gondor di rohan fratelli miei vedo nei vostri occhi la stessa paura che potrebbe afferrare il mio cuore ci sarà un giorno in cui il coraggio degli uomini cederà in cui abbandoneremo gli amici e spezzeremo ogni legame di fratellanza ma non è questo il giorno ci sarà l'ora dei lupi e degli scudi frantumati quando l'era degli uomini arriverà al crollo ma non è questo il giorno quest'oggi combattiamo per tutto cio che ritenete caro su questa bella terra v'invito a resistere uomini dell'ovest

Si tratta di un piccolo brano tratto da *Il signore degli anelli (Il ritorno del re)* di J.R.R. Tolkien. Ha bisogno di qualche intervento ancora, per esempio le maiuscole, ma il senso è chiaro. Il testo originale è questo:

> Figli di Gondor! Di Rohan! Fratelli miei! Vedo nei vostri occhi la stessa paura che potrebbe afferrare il mio cuore! ei sarà un giorno, in cui il coraggio degli uomini cederà, in cui abbandoneremo gli amici e spezzeremo ogni legame di fratellanza, ma non è questo il giorno! ei sarà l'ora dei lupi e degli scudi frantumati quando l'era degli uomini arriverà al crollo, ma non è questo il giorno! Quest'oggi combattiamo! Per tutto ciò che ritenete caro su questa bella terra, v'invito a resistere! uomini dell'Ovest!

La tabella usata per la cifratura, completa di tutte le corrispondenze, è mostrata qui di seguito. Da notare che non siamo in grado di determinare quali simboli corrispondano alle lettere *j*, *k*, *w*, *x*, *y* visto che non compaiono nel testo cifrato.

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

Cosa è necessario comunicare al destinatario, oltre al messaggio stesso naturalmente, perché riesca a decifrare correttamente e senza sforzo il testo cifrato.

Come in precedenza, il destinatario deve ricevere il messaggio cifrato e in qualche modo deve essere a conoscenza della chiave, sostanzialmente deve possedere la tabellina che abbiamo ricostruito qui sopra.

8. Cifratura con frase chiave

Abbiamo cifrato la frase che segue con il metodo della frase chiave, prova a decifrare il testo e a ricostruire l'alfabeto cifrato.

QRIFVFVCIBEVRUWURQPIGICCUFEIRQJUGVWUCIPUJJINVKPINIJIKEUKPILECEPIHQEFVJPIR-
GVOEVWURPQSVWESURPVKFEJEPVGIWWURPQJIKERNVRPJIRVFUJCUKPJIGUQVSEREUGVRRULURR-
QPJEPEJIOIBBEUJIOIBBUNDUKIRRVJEGUJUUDIRRVJEFJUKVECOQKPVFUJCUAUKPUNISFUKPJEKU-
KENVRPICIWUNNDEIFVFVCIBEVRUEJJENVRVKNELECUGIHQIRGVWEWURUCCIJSVREIUERQVWEWUR-
QPEFEQGEGEUNESECIFUJKVRUGUWVRVCICVJVAUCENEPIIUCUBUIJGLVQAAEUJ

La tabella di analisi per questo nuovo testo è riportata di seguito. Abbiamo colorato in grigio le caselle corrispondenti a simboli che hanno una frequenza pari a zero ed evidenziato in colore rosso simboli con le frequenze più basse mentre i simboli con le frequenze maggiori sono in verde.

Tabella di analisi

	A	B	C	D	E	F	G	H	I	J	K	L	M
#	4	7	18	3	40	13	13	2	37	26	15	4	0
%	1,11	1,95	5,01	0,84	11,14	3,62	3,62	0,56	10,31	7,24	4,18	1,11	0,00

	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
#	12	4	20	14	30	6	0	45	34	12	0	0	0
%	3,34	1,11	5,57	3,90	8,36	1,67	0,00	12,53	9,47	3,34	0,00	0,00	0,00

A questo puno il testo va analizzato, passo dopo passo, letto e riletto alla ricerca di indizi che possano portare a individuare le prime parole e via via brani sempre più lunghi di testo. Al termine di questo (non facile) lavoro si ottiene il testo in chiaro che è mostrato qui sotto.

Una popolazione venuta dalle pianure, dove la terra costa cara, si è stabilita qui, portando gioventù, movimento, spirito d'avventura. S'incontrano per le strade uomini e donne ben nutriti, ragazzi e ragazze che sanno ridere e hanno ripreso il gusto per le feste campestri. Se si conta la vecchia popolazione, irriconoscibile da quando vive nell'armonia, e i nuovi venuti, più di diecimila persone devono la loro felicità a Elezéard Bouffier.

Si tratta di un brano tratto da *L'uomo che piantava alberi* di J. Giono. La tabella usata per cifrarlo è questa:

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

costruita a partire dalla frase chiave *Il linguaggio della matematica si rivela irragionevolmente efficace.* Nella tabella le lettere derivate direttamente dalla frase chiave sono evidenziate in colore rosso, le restanti sono state aggiunte in ordine alfabetico a partire dalla lettera *F*, l'ultima a far parte della chiave.

9. Il metodo di Vigenére

Sul blog Research in Action è possibile utilizzare la libreria sviluppata con Blockly nel corso delle attività del progetto: http://researchinaction.it/myblockly/criptography.html
In particolare le funzioni *VigenereCifra* e *VigenereDecifra* sono realizzate implementando i due algoritmi descritti in queste pagine.
Blockly è un generatore di codice open source sostenuto da Google: https://developers.google.com/blockly/.

Scegli una frase chiave e cifra, con il metodo di Vigenère, la frase *Il denaro non cura la malattia, solo i sintomi*. **Successivamente decifra la frase.**

Innanzitutto, come di consueto, ripuliamo la frase da cifrare: *ildenarononcuralamalattiasoloisintomi*. Per conodità prendiamo la stessa frase chiave che abbiamo usato nell'esempio precedente: *Il linguaggio della matematica si rivela irragionevolmente efficace* e costruiamo la tabellina con il testo da decifrare nella seconda riga e la frase chiave nella prima. Ricordiamo che la tabella deve avere tante colonne quanti sono i caratteri del testo da cifrare e quindi, se necessario, la frase chiave va ripetuta anche più volte o troncata (come in questo caso) per adeguarla al testo in chiaro.

i	l	l	i	n	g	u	a	g	g	i	o	d	e	l	l	a	m	a	t	e	m	a	t	i	c	a	s	i	r	i	v	e	l	a	i	r
i	l	d	e	n	a	r	o	n	o	n	c	u	r	a	l	a	m	a	l	a	t	t	i	a	s	o	l	o	i	s	i	n	t	o	m	i

A questo punto procediamo come descritto. Per cifrare la lettera *i* dobbiamo usare l'alfabeto che si trova nell'ottava riga della tabella di Vigenère (cfr. Metodo di Vigenère a pagina xx) e scegliere il simbolo che si trova nella colonna corrispondente alla lettera da cifrare (la colonna con la *i*) per cui il simbolo da usare è *Q* (la riga usata è evidenziata in colore verde e il simbolo in rosso). Non è mostrata qui l'intera tabella ma, per conodità di lettura, solo alcune righe.

Il secondo carattere è *l*, la lettera corrispondente nella frase chiave è anch'essa una *elle* per cui dobbiamo far riferimento all'undicesima riga (anche questa evidenziata in colore verde) e scegliere il simbolo corrispondente alla colonna che nella prima riga ha la *elle*: *W*.

Ancora, la terza lettera è d, a cui corrisponde nella frase chiave ancora una elle ma questa volta dobbiamo scegliere il simbolo che si trova nella colonna corrispondente alla lettera d dell'alfabeto in chiaro: O.

Proseguendo in questo modo otteniamo *QWOMAGLOTUVQXVLWAYAEEFTBIUODWZADREOUZ*.

Tabella di Vigenère

	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
...																										
6	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
7	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	G
8	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
9	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	I
10	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
11	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
12	M	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
13	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
...																										

Per la decifrazione si procede in modo simile. Prima di tutto costruiamo la tabella per decifrare il messaggio appaiando testo cifrato e frase chiave, come mostrato qui di seguito.

i	l	l	i	n	g	u	a	g	g	i	o	d	e	l	l	a	m	a	t	e	m	a	t	i	c	a	s	i	r	i	v	e	l	a	i	r
Q	W	O	M	A	G	L	O	T	U	V	Q	X	V	L	W	A	Y	A	E	E	F	T	B	I	U	O	D	W	Z	A	D	R	E	O	U	Z

Il primo simbolo è *Q* che corrisponde a i nella frase chiave: dobbiamo usare l'ottava riga e cercare proprio la *Q* che si trova nella colonna che in testa ha la lettera *i*. Procedendo in questo modo si decifra l'intero testo, si tratta in sostanza dello stesso meccanismo usato per la cifratura per così dire *al contrario*.

Prova a determinare un algoritmo, una procedura, per cifrare un testo usando il metodo di Vigenère senza far riferimento alla tabella omonima.

Riprendiamo l'esempio precedente. Per cifrare la prima lettera abbiamo scelto il simbolo che si trova nell'ottava (8) riga e nella nona (9) colonna, questo simbolo è la Q che à la diciassetesima (17) lettera dell'alfabeto: $8 + 9 = 17$. Notiamo anche che la riga da usare corrisponde alla posizione della lettera della chiave nell'alfabeto, diminuita di uno (infatti la *i* è la 9° lettera dell'alfabeto e abbiamo usato la riga numero 8).

Per cifrare la terza lettera abbiamo scelto il simbolo nell'undicesima (11) riga e quarta (4) colonna, si tratta della *O* che è la 15-esima lettera dell'alfabeto e, ancora una volta, $11 + 4 = 15$! Anche questa volta abbiamo che la riga da usare si ottiene sottraendo uno alla posizione della lettera della chiave (la *elle* è la 12° lettera e la riga usata è la numero 11).

C'è un legame tra 18 e 3 e 21? Questa relazione vale per ogni altro simbolo del testo cifrato?

Possiamo azzardare un'ipotesi: sommando la riga e la colonna abbiamo la posizione del simbolo cifrante nell'alfabeto!

Proviamo con la quart'ultima lettera: la t. Dobbiamo usare ancora la riga corrispondente alla l, l'undicesima (11) e scegliere il simbolo che si trova nella 20-esima colonna: la E. In questo caso 11 + 20 = 31 ma la E è al quinto posto nell'alfabeto. Non funziona?

Beh! In realtà sì se torniamo a ragionare in termini di modulo: la posizione nell'alfabeto deve essere un numero intero compreso tra 1 e 26 e $31 \bmod 26 = 5$, proprio la posizione del simbolo *E*!

Determina un algoritmo, una procedura, per decifrare un testo che è stato cifrato con il metodo di Vigenère e come in precedenza, fai in modo che l'algoritmo non abbia bisogno della tabella.

Quindi la procedura, formalizzata in un italiano abbastanza preciso, potrebbe essere come questa descritta qui di seguito.

sia i la posizione del carattere nel testo in chiaro e c_chiaro il carattere stesso (quello da cifrare)

```
c_chiaro il carattere da cifrare
i posizione del carattere stesso nel testo in chiaro
determinare la lettera della chiave corrispondente alla lettera
      n = i mod lunghezza della chiave
      c_chiave = lettera nella chiave che si trova nella posizione n
determinare la riga corrispondente nella tabella di Vigenère
che corrisponde alla posizione di c_chiave nell'alfabeto
diminuita di uno
      y = la posizione di c_chiave nell'alfabeto - 1
determinare la posizione nell'alfabeto del carattere in chiaro
      z = la posizione di c_inchiaro nell'alfabeto
determinare la posizione del simbolo corrispondente al carattere
in chiaro nel messaggio cifrato
      p = (y+z) mod lunghezza dell'alfabeto
      c_cifrato = carattere nell'alfabeto la cui posizione è p
```

Per decifrare un simbolo possiamo usare, come già anticipato, il procedimento inverso. Se sommando riga e colonna abbiamo la posizione del simbolo nell'alfabeto allora la colonna (e quindi, indirettamente, la lettera in chiaro) si può ottenere sottraendo alla posizione del simbolo la riga

nell'alfabeto! Nel dettaglio:

```
c_cifrato il carattere del messaggio cifrato
i posizione del carattere stesso nel testo cifrato
determinare la lettera della chiave corrispondente al simbolo cifrato
      n = i mod lunghezza della chiave
      c_chiave = lettera nella chiave che si trova nella posizione n
determinare la riga corrispondente nella tabella di Vigenère
che corrisponde alla posizione di c_chiave nell'alfabeto
diminuita di uno
      y = la posizione di c_chiave nell'alfabeto - 1
determinare la posizione nell'alfabeto del carattere in chiaro
      z = la posizione di c_cifrato nell'alfabeto
determinare la posizione del simbolo corrispondente al carattere
in chiaro nel messaggio cifrato
      p = (z-y) mod lunghezza dell'alfabeto
      c_inchiaro = carattere nell'alfabeto la cui posizione è p
```

Sul blog Research in Action è possibile utilizzare la libreria sviluppata con Blockly nel corso delle attività del progetto: http://researchinaction.it/myblockly/criptography.html
In particolare le funzioni *VigenereCifra* e *VigenereDecifra* sono realizzate implementando i due algoritmi descritti in queste pagine.
Blockly è un generatore di codice open source sostenuto da Google: https://developers.google.com/blockly/.

9.1. La procedura in Python

Per curiosità, ma anche per completezza, presentiamo qui di seguito la procedura così come ottenuta attraverso il generatore di codice Blockly.

```
def VigenereCifra(testoinchiaro, chiave):
      # le istruzioni per preparare il testo e la chiave sono omesse

      for i in range(1, len(testoinchiaro):
            # Determina la posizione della lettera corrispondente
            # nella frase chiave
            if i % len(chiave) == 0:
                  n = len(chiave)
            else:
                  n = i % len(chiave)
            c_chiave = chiave[int(n - 1)]

            # Determina la riga da usare nella tabella di Vigenère
            if c_chiave == 'a':
                  y = 26
            else:
                  y = first_index(alfabetoinchiaro, c_chiave) - 1
            c_inchiaro = testoinchiaro[i - 1]
            z = first_index(alfabetoinchiaro, c_inchiaro)

            # Determina la posizione del simbolo cifrato come somma
            # di riga e colonna modulo lunghezza dell'alfabeto
            if (y + z) % len(alfabetoinchiaro) == 0:
                  n = len(alfabetoinchiaro)
            else:
                  n = (y + z) % len(alfabetoinchiaro)
            testocifrato = str(testocifrato) + str(alfabetoinchiaro[n - 1])
      return testocifrato.upper()
```

10. Struttura della libreria

I paragrafi qui di seguito riportano le dipendenze tra le varie funzioni della libreria dedicate all'approssimazione dei numeri irrazionali e trascendenti: le funzioni più in alto nella figura usano le funzioni più in basso a cui sono collegate. Per esempio: *CodiceCesareCifra* usa *PreparaTestoPer-Cifratura* che a sua volta usa *TrasformaAccentate*.

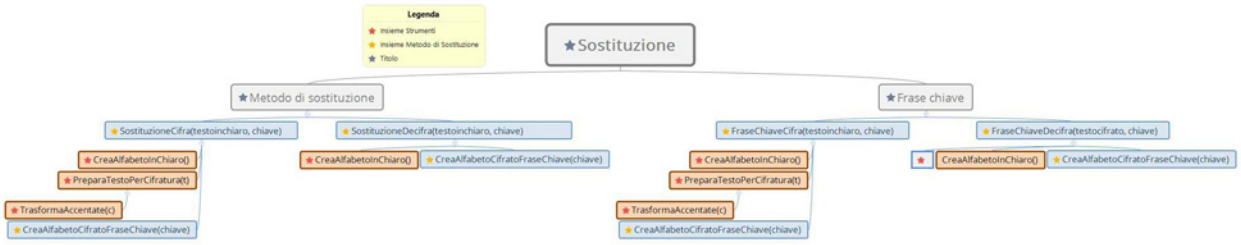
La legenda indica in quale insieme di funzioni possono essere trovate le procedure corrispondenti (le piccole stelle colorate presenti davanti al nome di ogni funzione son riportate nella legenda, in cui è indicato l'insieme di funzioni opportuno).

10.1. Metodo di Cesare

Funzioni per cifrare e decifrare con il metodo di Cesare **(cfr. 2 Metodo di Cesare a pagina 7)**.

10.2. Metodo di sostituzione

Funzioni per cifrare e decifrare con il metodo di sostituzione o con il metodo della frase chiave (cfr. **Cifratura per sostituzione a pagina 10** e **Cifratura con frase chiave a pagina 14**).



10.3. Metodo di Vigenere

Funzioni per cifrare e decifrare con il metodo di Vigenère **(cfr. 5 Metodo di Vigenère a pag. 16)**.

LICEO SCIENTIFICO GRASSI LATINA

CNR IAC
Istituto per le Applicazioni del Calcolo

CNR IFN
Istituto di Fotonica e Nanotecnologie

INM

ISMAR
Istituto di Scienze Marine

LSS G.B. Grassi

LICEO SCIENTIFICO STATALE G.B. GRASSI DI LATINA

WWW.LICEOGRASSILATINA.ORG

CNR - IAC

ISTITUTO PER LE APPLICAZIONI DEL CALCOLO MAURO PICONE

WWW.IAC.CNR.ORG

CNR - IFN Roma

ISTITUTO DI FOTONICA E NANOTECNOLOGIE

WWW.ROMA.IFN.CNR.ORG

CNR - INM

ISTITUTO DI INGEGNERIA DEL MARE

WWW.INSEAN.CNR.ORG

CNR - ISMAR

ISTITUTO DI SCIENZE MARINE

WWW.ISMAR.CNR.ORG