



PEGASO
Università Telematica



Algebra Modulare e Introduzione alla Crittografia

Corso di Algebra

Laurea in Informatica

In questa lezione riprenderemo i concetti delle classi di resto e della divisibilità, per trattare più nel dettaglio queste tematiche dal punto di vista algebrico, e per parlarne dal punto di vista applicativo, con un accenno anche a tematiche di crittografia.

1 Aritmetica modulare

Con aritmetica modulare si intende l'insieme delle operazioni aritmetiche standard, ovvero somma, sottrazione, moltiplicazione e divisione, e di studiarle nel contesto delle classi di resto modulo n . A tale scopo, richiamiamo le definizioni principali:

Definizione 1.1. Sia $(\mathbb{Z}, +, \cdot)$ l'anello dei numeri interi. Dato un numero intero n ed un numero intero a tale che $0 \leq a \leq n - 1$, la **classe di resto a modulo n** , indicata con \bar{a} , è costituita da tutti i numeri interi che, divisi per n , diano resto pari ad a .

La classe di resto $\bar{0}$ è un sottogruppo di $(\mathbb{Z}, +)$, mentre le classi di resto \bar{a} con $1 \leq a \leq n - 1$ sono i suoi laterali.

L'insieme delle classi di resto $\mathbb{Z}/n\mathbb{Z} = \{\bar{0}, \bar{1}, \dots, \overline{n-1}\}$ costituiscono un **anello** rispetto alla somma e alla moltiplicazione.

Inoltre, se n è primo, allora $\mathbb{Z}/n\mathbb{Z}$ è anche un **campo**, ovvero ogni elemento $a \neq \bar{0}$ è invertibile (rispetto alla moltiplicazione).

Per quanto riguarda l'aritmetica modulare, per fortuna le operazioni sono abbastanza semplici da fare. Infatti:

Osservazione 1.2. In $\mathbb{Z}/n\mathbb{Z}$ valgono le seguenti proprietà:

- $\bar{a} + \bar{b} = \overline{a + b}$
- $\bar{a} - \bar{b} = \overline{a - b}$
- $\bar{a} \cdot \bar{b} = \overline{a \cdot b}$
- $\bar{a}^{xy} = (\bar{a}^x)^y$

ovvero, se dobbiamo fare la somma/differenza/moltiplicazione tra due numeri modulo n , possiamo

1. calcolare la loro somma/differenza/moltiplicazione come se fossero numeri interi normali;
2. calcolare il resto della divisione del risultato modulo n .

Esempio 1.3. Consideriamo $n = 35$, e supponiamo di voler calcolare $25 + 30$, $20 - 24$ e $20 \cdot 3$ modulo n . Allora:

1. Calcoliamo ciascuno di questi numeri come se fossero in \mathbb{Z} :

$$25 + 30 = 55$$

$$20 - 24 = -4$$

$$20 \cdot 3 = 60$$

2. Valutiamo il resto della divisione del risultato per $n = 35$:

$$55 = 35 + \mathbf{20}$$

$$-4 = -35 + \mathbf{31}$$

$$60 = 35 + \mathbf{25}$$

3. Quindi, abbiamo che:

$$25 + 30 = 20 \pmod{35}$$

$$20 - 24 = 31 \pmod{35}$$

$$20 \cdot 3 = 25 \pmod{35}$$

Dove di solito il $\pmod{35}$ si indica per ricordarsi che non stiamo lavorando nei numeri interi. Si tenga presente che spesso si tende ad omettere se è chiaro dal contesto.

Esempio 1.4. Consideriamo $n = 8$, e costruiamo

$$\mathbb{Z}/8\mathbb{Z} = \{\bar{0}, \bar{1}, \bar{2}, \bar{3}, \bar{4}, \bar{5}, \bar{6}, \bar{7}\}$$

Abbiamo che $(\mathbb{Z}/8\mathbb{Z}, +, \cdot)$ é un anello; possiamo quindi costruire la tabella relativa alla somma e alla moltiplicazione. Per semplicità, omettiamo la scrittura della classe di resto, e scriviamo solo a anzichè \bar{a} :

+	0	1	2	3	4	5	6	7	·	0	1	2	3	4	5	6	7
0	0	1	2	3	4	5	6	7	0	0	0	0	0	0	0	0	0
1	1	2	3	4	5	6	7	0	1	0	1	2	3	4	5	6	7
2	2	3	4	5	6	7	0	1	2	0	2	4	6	0	2	4	6
3	3	4	5	6	7	0	1	2	3	0	3	6	1	4	7	2	5
4	4	5	6	7	0	1	2	3	4	0	4	0	4	0	4	0	4
5	5	6	7	0	1	2	3	4	5	0	5	2	7	4	1	6	3
6	6	7	0	1	2	3	4	5	6	0	6	4	2	0	6	4	2
7	7	0	1	2	3	4	5	6	7	0	7	6	5	4	3	2	1

Ad esempio, leggiamo che $3 \cdot 5 = 7$, poiché $3 \cdot 5 = 15$ e $15 = 8 + 7$, quindi $3 \cdot 5 = 7 \pmod{8}$.

La tabella moltiplicativa ci permette anche di evidenziare che **non sempre** un elemento ha un inverso moltiplicativo, ovvero dato a non sempre esiste b tale che $a \cdot b = 1$; ad esempio, $3 \cdot 3 = 1$ (quindi, 3 ammette inverso, ed è proprio se stesso), ma 2 non ammette nessun inverso.

Abbiamo visto inoltre che:

Osservazione 1.5. Il gruppo $(\mathbb{Z}/n\mathbb{Z}, +)$ è ciclico, e un elemento $a \in \mathbb{Z}/n\mathbb{Z}$ è un generatore di $\mathbb{Z}/n\mathbb{Z}$ se e soltanto se $MCD(a, n) = 1$.

Esempio 1.6. Considerando sempre $\mathbb{Z}/8\mathbb{Z}$, abbiamo che 3 è un generatore del gruppo. Infatti abbiamo che, modulo 8:

$$\begin{aligned} 3 + 3 &= 6 & 9 + 3 &= 9 & 9 + 3 &= 12 = 4 \\ 4 + 3 &= 7 & 7 + 3 &= 10 = 2 & 2 + 3 &= 5 \\ 5 + 3 &= 8 = 0 \end{aligned}$$

Alternativamente, si può vedere che nella riga (o colonna) della tabella moltiplicativa relativa

all'elemento 3 troviamo tutti gli elementi del gruppo.

2 Calcolo dell'inverso modulo n

Dato un elemento $a \in \mathbb{Z}/n\mathbb{Z}$, possiamo calcolare l'elemento inverso modulo n utilizzando l'algoritmo di Euclide.

Allora, per trovare l'elemento inverso di un elemento a dato in modulo n , possiamo utilizzare l'algoritmo di Euclide, in cui dobbiamo iniziare con il resto della divisione tra n e a . Quindi:

$$\begin{aligned} n &= q_0 a + r_0 \\ a &= q_1 r_0 + r_1 \\ r_0 &= q_2 r_1 + r_2 \\ r_1 &= q_3 r_2 + r_3 \\ &\vdots \\ r_{n-2} &= q_n r_{n-1} + r_n \\ r_{n-1} &= q_{n+1} r_n + 1 \end{aligned}$$

Dopo un certo numero di iterazioni dell'algoritmo, si arriverà alla fine **necessariamente** ad avere resto pari a 1. Quando questo succede, si può iniziare tramite sostituzione all'indietro a risalire le iterazioni dell'algoritmo; iniziamo ad invertire le relazioni trovate precedentemente:

$$\begin{aligned} 1 &= r_{n-1} - q_{n+1} r_n \\ r_n &= r_{n-2} - q_n r_{n-1} \\ &\vdots \\ r_3 &= r_1 - q_3 r_2 \\ r_2 &= q_2 r_1 - r_0 \\ r_1 &= q_1 r_0 - a \\ r_0 &= q_0 a - n \end{aligned}$$

ed iniziamo a sostituire all'indietro. Ad esempio (qui mostriamo solo le prime iterazioni, nell'esempio

che segue sarà più chiaro)

$$r_1 = q_1 r_0 - a = q_1(q_0 a - n) - a$$

$$r_2 = q_2 r_1 - r_0 = q_2(q_1(q_0 a - n) - a) - (q_0 a - n)$$

e così via. Si arriverà in fondo ad una relazione del tipo:

$$1 = a \cdot b \pmod{n}$$

Da cui quindi possiamo vedere il valore dell'inverso di a .

Esempio 2.1. Consideriamo $n = 3120$, e supponiamo di voler trovare il valore dell'inverso di $a = 17$. Iniziamo ora l'algoritmo di Euclide con $n = 3120$ e $a = 17$:

$$3120 = 183 \cdot 17 + 9$$

$$17 = 1 \cdot 9 + 8$$

$$9 = 1 \cdot 8 + 1$$

Ci interrompiamo quindi a questo step, dato che abbiamo ottenuto resto 1. Ora, invertiamo le relazioni precedenti, mettendo in evidenza il resto di ogni step:

$$9 = 3120 - 183 \cdot 17$$

$$8 = 17 - 1 \cdot 9$$

$$1 = 9 - 1 \cdot 8$$

Adesso, sostituiamo il valore 8 dell'ultima relazione con $17 - 1 \cdot 9$ (lasciando in evidenza il termine 9, che sarà il prossimo resto da sostituire):

$$1 = 9 - 1 \cdot 8 = 9 - 1 \cdot (17 - 1 \cdot 9) = 2 \cdot 9 - 17$$

Sostituiamo quindi il valore 9 con quanto scritto nella prima relazione, lasciando in evidenza il valore di $a = 17$:

$$\begin{aligned} 1 &= 2 \cdot 9 - 17 = 2 \cdot (3120 - 183 \cdot 17) - 17 \\ &= 2 \cdot 3120 - 366 \cdot 17 - 17 \\ &= 2 \cdot 3120 - 367 \cdot 17 \end{aligned}$$

Quindi, rileggendo l'ultima espressione ottenuta, modulo 3120, otteniamo:

$$1 = -367 \cdot 17 \pmod{3120}$$

Quindi l'inverso di $a = 17$ modulo $n = 3120$ è pari a $-367 \pmod{3120} = 2753 \pmod{3120}$.

Chiaramente, per esempi con valori di n piccoli può essere più comodo cercare l'inverso "a mano"; tuttavia, si vede facilmente il vantaggio di utilizzare questo algoritmo per valori grandi di n .

3 Metodo di fattorizzazione di Fermat

In questa sezione esamineremo un metodo di fattorizzazione dei numeri naturali dispari, proposta da Fermat.

Dato un numero intero dispari n , sia $a = \lfloor \sqrt{n} \rfloor$, ovvero il numero intero appena più grande della radice quadrata di n (ad esempio, $\sqrt{17} \approx 4.12310\dots$, quindi $\lfloor \sqrt{17} \rfloor = 4$). Ripetiamo quindi i seguenti passaggi:

1. Definiamo $b' = a^2 - n$;
2. Se b' non è un quadrato perfetto, ripeti il passaggio precedente con $a = a + 1$;

Quindi, il numero b' , usciti da questa iterazione, è un quadrato perfetto. Definiamo $b = \sqrt{b'}$.

Allora, una fattorizzazione di n è:

$$n = (a - b)(a + b)$$

Esempio 3.1. Prendiamo come esempio il numero dispari $n = 357$. Abbiamo che $\sqrt{357} \approx 18.574\dots$, quindi iniziamo l'algoritmo con $a = 19$.

Abbiamo che $19^2 - 357 = 4$, che è un quadrato perfetto, quindi $b = \sqrt{4} = 2$ e una fattorizzazione di n è data da:

$$357 = (19 - 2) \cdot (19 + 2) = 17 \cdot 21$$

Esempio 3.2. Consideriamo $n = 155$. Abbiamo che $\sqrt{155} \approx 12.450\dots$, quindi iniziamo l'algoritmo con $a = 13$.

Abbiamo che:

a	13	14	15	16	17	18
$a^2 - n$	14	41	70	101	134	169
è un quadrato perfetto?	No	No	No	No	No	Si

Quindi $b = \sqrt{169} = 13$ e abbiamo:

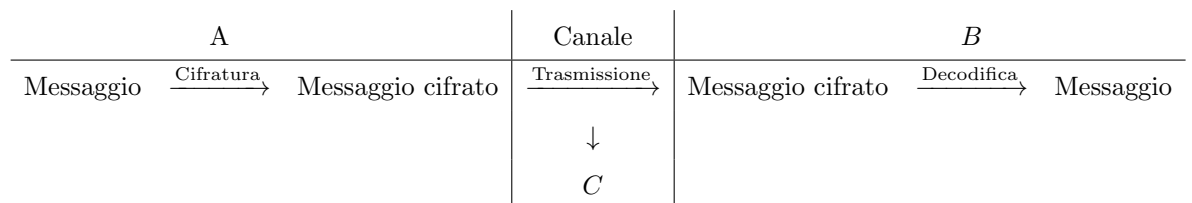
$$n = (18 - 13) \cdot (18 + 13) = 5 \cdot 31$$

4 Introduzione alla crittografia

Definiamo prima concettualmente cosa sia e a cosa serva la crittografia:

Ci sono due interlocutori, Alice (A) e Bob (B) che vogliono inviarsi un messaggio. Tuttavia, questo messaggio contiene informazioni riservate tra loro due, e pertanto non vogliono che una eventuale terza persona Charlie (C) possa leggere il messaggio. Il modo più semplice per far sì che il messaggio non venga letto da un eventuale C consiste nel trovarsi di persona per consegnare il messaggio. Tuttavia, questo chiaramente non è sempre possibile, né risulta essere l'opzione più pratica.

A tale scopo, il messaggio viene trasmesso o inviato da A a B per mezzo di un canale. Non è assicurato che C non vi possa accedere a questo canale, pertanto quello che A e B fanno è concordarsi su quella che è la cifratura del messaggio; ovvero, il messaggio viene trasformato da A da un testo leggibile (detto **in chiaro**) in un modo che sia illeggibile per chiunque intercetti il messaggio, e che sia trasformabile nel messaggio originale solo da B:



Una eventuale terza parte C ha solo accesso al messaggio trasmesso, che è cifrato. Pertanto, a meno che C non conosca il modo per decifrare il codice e recuperare il

messaggio originale, la trasmissione tra A e B è sicura.

4.1 Cifrari a sostituzione

Vediamo ora un esempio di messaggio crittografato molto semplice, chiamato Cifrario di Cesare:

Definizione 4.1. Dato un messaggio in lettere latine ABCD...Z, la sua cifratura ottenuta con il **Cifrario di Cesare** si ottiene sostituendo ogni lettera del messaggio con la lettera situata ad un certo numero di posizioni dopo nell'alfabeto.

Esempio 4.2. Il messaggio:

HELLO WORLD

Cifrato secondo il Cifrario di Cesare, in cui si sceglie la lettera posizionata 2 posizioni dopo, è:

JGNNQ YQTNF

Per decodificare il messaggio, quindi, bisogna spostare all'indietro ogni lettera di due posizioni.

Questo cifrario, che tradizione vuole essere stato sviluppato proprio da Giulio Cesare, trova ben poche speranze di rimanere sicuro al giorno d'oggi. Innanzitutto, è stato sviluppato in un'epoca nella quale la maggior parte delle persone non sapeva leggere un testo qualunque, figuriamoci uno cifrato. Poi, con l'aumentare della scolarizzazione e del numero di persone in grado di leggere, ci si è resi conto che questo genere di crittografia non è propriamente sicura.

Si potrebbe dire che un'alternativa valida consiste nel sostituire una lettera con una casuale dell'alfabeto. (Ad esempio, tutte le lettere C diventano G, le A diventano K, eccetera, senza nessuna relazione tra di esse)

Questo porterebbe il numero di combinazioni possibili a $26!$ lettere, che è un numero tutto sommato sufficientemente grande da risultare una sfida anche per i moderni calcolatori. Tuttavia, la grande debolezza di questo tipo di cifrari è dovuto al fatto che rimangono completamente attaccabili nel caso di messaggi sufficientemente lunghi (anche poche righe). Prendiamo ad esempio questo messaggio cifrato:

IPVQCQCIQGKSFVPPQFDFQCKLKQNNPSQLQVLBV
FVVHVYKDLVZVSSKNNPVFVVEZVYIQKMQNPQIVFQ
VCVSRKFIQLBKNEPVGPIPIPVQVEPIK

Notiamo innanzitutto che il messaggio è apparentemente incomprensibile. Per buona misura, nei messaggi trasmessi si tolgono anche gli spazi tra le parole, rendendo ancora più complessa l'analisi del messaggio da parte di terzi.

Immaginiamo di essere C, e di voler decodificare questo messaggio. Innanzitutto, non possiamo sperare di poter recuperare il messaggio dandolo al computer e facendogli provare tutte le combinazioni, dato che

$$26! = 403.291.461.126.605.635.584.000.000$$

In generale, allo stato attuale dei calcolatori si può considerare come limite di computazione circa 10^{14} operazioni, ovvero:

$$100.000.000.000.000$$

Dobbiamo quindi cercare un'altra strada. Iniziamo con introdurre quindi l'**analisi delle frequenze** del messaggio, ovvero contiamo quante volte appare una lettera nel messaggio, iniziando da quelle più frequenti; teniamo conto del fatto che il messaggio ha 105 lettere. Ad esempio:

- La lettera V appare 17 volte, ovvero circa il 18%;
- La lettera Q appare 14 volte, ovvero circa il 13%;
- La lettera P appare 11 volte, ovvero circa il 10%;

E così via. Abbiamo quindi la seguente tabella:

V	Q	P	K	F	I	N	L	S	C	E	B	D	G	Y	Z	H	M	R
19	14	11	9	8	8	6	5	5	4	3	2	2	2	2	2	1	1	1

Dato che gli interlocutori sono italiani, possiamo presumere che anche il messaggio sia in italiano. L'appiglio, quindi, nella fase in cui cerchiamo di **forzare il codice**, sta nel fatto che le lettere nelle parole non hanno tutte la stessa frequenza. In particolare, le **vocali** sono le lettere che appaiono più spesso (tranne la U, che è relativamente più rara). Possiamo quindi immaginare quattro lettere tra V, Q, P, K e N corrispondano alle vocali A, E, I, O.

Vediamo inoltre che il messaggio inizia con IPVQCQC e finisce con IPVQVEPIK ed essendo un messaggio tra due amici è probabile che all'inizio e alla fine del messaggio vi siano dei convenevoli. In particolare, notiamo che abbiamo la stessa parte iniziale, IPVQ, che contiene ben

3 delle lettere candidate ad essere delle vocali. Possiamo quindi ipotizzare che $IPVQ = CIAO$, e che $P = I$, $V = A$ e $Q = O$. È ragionevole pensare, inoltre, che $CQC = BOB$, e $VEPIK = ALICE$, quindi $C = B$, $E = L$, $I = C$ e $K = E$. Sostituiamo queste ipotesi nel messaggio originale (le evidenziamo in grassetto), e aggiungiamo qualche spazio per dare un'idea migliore di come sia strutturato il messaggio (teniamo presente, ad esempio, che la stessa vocale non appare due volte di fila nella stessa parola):

**CIAO BOB COGESFAI IOFDFFOBELEONNISOLOALBA
FA AHAYEDLAZASSENNIAFA ALZAYCOEMONIOCAFO
ABASREFCOLBENLIAGICI CIAO ALICE**

In fondo al messaggio, vediamo le lettere **LIAGICI**. Possiamo supporre ci sia uno spazio tra **LI** e **AGICI**, e che quindi l'ultima parola sia "AMICI", quindi $G = M$; se questo non fosse il caso, possiamo sempre tornare indietro. Sotto questa ipotesi, abbiamo all'inizio la frase **COMESFAI**, che potrebbe essere "COME STAI", quindi $S = S$ e $F = T$, e inoltre la parola precedente ad amici potrebbe essere **GLI**, quindi $N = G$:

**CIAO BOB COME STAI IOTDTTTOBELE OGGI SOLOALBA
TA AHAYEDLAZASSEGGIATA ALZAYCOEMO GIOCATO
ABASRETCOLBE GLIAMICI CIAO ALICE**

Probabilmente, $Z = P$ e la parola è PASSEGGIATA, e **COEMO GIOCATO** potrebbe essere costituito dalle parole **CO E MO GIOCATO**. All'inizio del messaggio, **IOTDTTTOBELE** potrebbe essere **IO TUTTO BENE**, da cui $D = U$ e $L = N$:

**CIAO BOB COME STAI IOTUTTOBENE OGGI SONOANBA
TA AHAYEUNAPASSEGGIATA ALPAYCO E MO GIOCATO
ABASRETCONBE GLIAMICI CIAO ALICE**

È quindi semplice recuperare il messaggio originale:

**CIAO BOB COME STAI IO TUTTO BENE OGGI SONO ANDA
TA A FARE UNA PASSEGGIATA AL PARCO E HO GIOCATO
A BASKET CON DEGLI AMICI CIAO ALICE**

Vediamo che facendo delle ipotesi sulla frequenza delle vocali, quindi, siamo riusciti a ricostruire agevolmente il messaggio; se in alcune parti del messaggio abbiamo dovuto azzardare delle ipotesi per andare avanti, tutto sommato queste sono state confermate da evidenze presenti in altri punti del messaggio, e non abbiamo ottenuto parti prive di senso. Inoltre, bisogna immaginare che se noi siamo riusciti a decifrare a mano un messaggio di questo tipo, un calcolatore odierno programmato correttamente sarebbe in grado di decifrare questo messaggio in pochi millisecondi; se ne deduce che questo tipo di cifrari possono essere ritenuti abbastanza superati, a meno di inventare anche una vera e propria lingua da zero in cui scrivere i messaggi.

Sebbene la presenza dei convenevoli iniziali e finali ci abbia aiutato non poco nella ricostruzione del messaggio, questo non è da pensare come una cosa irrealistica. Spesso, i sistemi di crittografia sarebbero anche sicuri, ma l'operatore umano che deve adottarli risulta "pigro" nell'essere sicuro di utilizzarli al meglio. Ad esempio, durante la Seconda Guerra Mondiale i tedeschi dell'Asse avevano sviluppato una macchina (Enigma) praticamente impossibile da decifrare con i mezzi dell'epoca (fatto salvo sapere esattamente come era costruita la macchina). Ciononostante, un grande appiglio dato agli Alleati per capirne il funzionamento era dato dall'abitudine di utilizzare frasi ricorrenti nei messaggi, come Wetterbericht (bollettino meteo), An alle Einheiten (a tutte le unità), che pur essendo cifrate erano riconoscibili all'inizio del testo.

Nella prossima lezione vedremo meglio come l'algebra può venire in aiuto per la definizione di codici sicuri.

Test di autovalutazione

Si risponda alle seguenti domande sulla presente lezione. Ogni domanda ammette esattamente una risposta corretta.

1. In \mathbb{Z}_n , due numeri a e b sono congruenti modulo n se:
 - (A) $a - b$ è multiplo di n
 - (B) $a + b$ è multiplo di n
 - (C) a e b sono entrambi primi
 - (D) n è un numero primo
2. In \mathbb{Z}_n , la somma modulo n di $\overline{7}$ e $\overline{5}$ in \mathbb{Z}_{10} è:
 - (A) $\overline{12}$
 - (B) $\overline{2}$
 - (C) $\overline{3}$
 - (D) $\overline{5}$
3. L'elemento $\overline{a} \in \mathbb{Z}_n$ è invertibile se e solo se:
 - (A) a è primo
 - (B) $a < n$
 - (C) $\gcd(a, n) = 1$
 - (D) n è dispari
4. Calcolare l'inverso di 3 in \mathbb{Z}_7 :
 - (A) l'inverso è 2
 - (B) l'inverso è 5
 - (C) l'inverso è 4
 - (D) l'inverso è 3
5. Il metodo di fattorizzazione di Fermat si basa su:
 - (A) scrivere un numero come differenza di due quadrati

- (B) scrivere un numero come somma di due quadrati
 - (C) calcolare il massimo comune divisore
 - (D) trovare l'inverso modulo n
6. Usando il metodo di Fermat, fattorizzare $n = 595$:
- (A) abbiamo $n = 5 \cdot 119$
 - (B) abbiamo $n = 35 \cdot 17$
 - (C) abbiamo $n = 29 \cdot 21$
 - (D) abbiamo $n = 13 \cdot 47$
7. Quanto vale $17^{12} \pmod{51}$:
- (A) vale 17
 - (B) vale 1
 - (C) vale 34
 - (D) non è calcolabile esattamente, essendo un numero troppo grande
8. Nel cifrario a sostituzione semplice:
- (A) ogni lettera del testo in chiaro è sostituita sempre dalla stessa lettera cifrata
 - (B) ogni lettera è sostituita da una cifra casuale diversa ogni volta
 - (C) il testo cifrato è sempre più lungo del testo in chiaro
 - (D) si usa una matrice di permutazione
9. L'analisi delle frequenze sfrutta:
- (A) la distribuzione uniforme delle lettere
 - (B) le diverse frequenze delle lettere
 - (C) la lunghezza fissa delle parole
 - (D) la conoscenza della chiave
10. Quale tra questi **non** è una strategia efficace per cercare di decifrare un messaggio cifrato con il metodo di sostituzione:

- (A) provare le combinazioni possibili
- (B) cercare delle espressioni di uso comune (saluti, firma, ecc)
- (C) cercare di identificare quali lettere possano essere delle vocali
- (D) ipotizzare la lingua in cui è scritto il messaggio