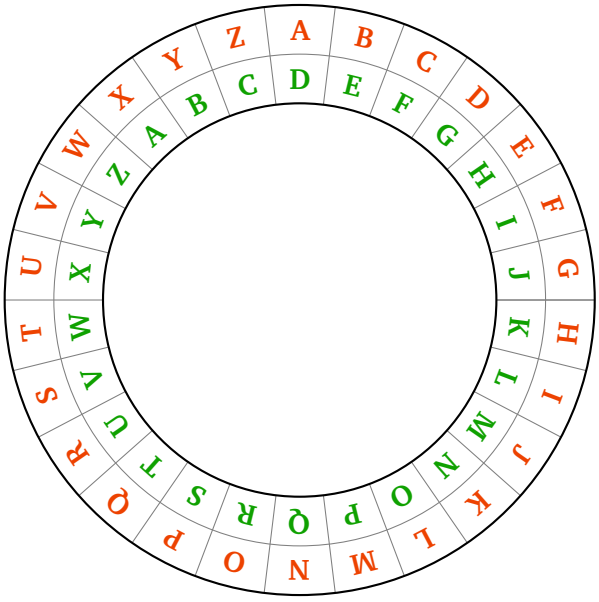
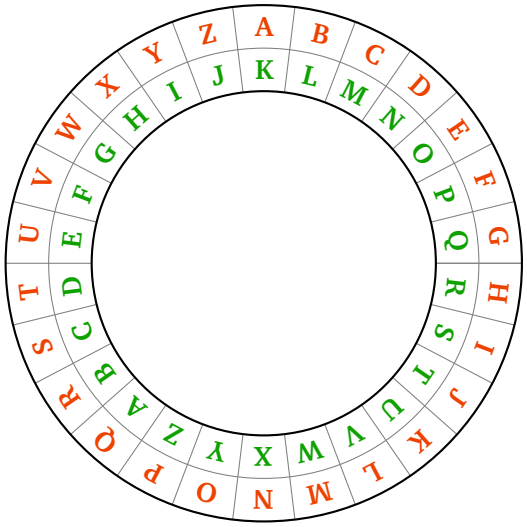
Cryptographie

Activité 1 (Le code de César)**.**
Jules César transmettait ses messages de façon cachée. Par exemple, le message **ALLEZ ASTERIX** était transformé en **DOOHC DVWHULA**. Chaque lettre était décalée dans l'ordre alphabétique de 3 lettres : **A** devenait **D**, **B** devenait **E**, **C** devenait **F**... Lorsque l'on arrivait à la fin de l'alphabet (**W** devenait **Z**), on repartait du début : **X** devenait **A**, **Y** devenait **B** et **Z** devenait **C**. La roue ci-dessous t'aide pour coder le message, une lettre sur l'anneau extérieur est codée en la lettre en vis à vis de l'anneau intérieur.

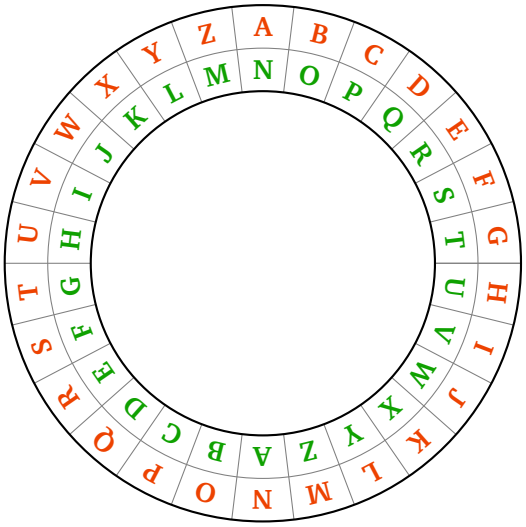


Pour décoder un message, on recule de trois lettres : **D** se décode en **A**, **E** se décode en **B**... Autrement dit, on passe de l'anneau intérieur à l'anneau extérieur.

1. (a) Code la phrase **ABC DE L INFO DEPUIS ZERO**.
   (b) Trouve le nom de la première femme informaticienne en décodant **DGD ORYHODFH**.
2. Bien sûr, on peut changer le décalage. Avec un décalage de 10 : **A** devient **K**, **B** devient **L**... (roue de gauche). Avec un décalage de 13 : **A** devient **N**, **B** devient **O**... (roue de droite).



Décalage de 10



Décalage de 13

(a) Pour un décalage de 10 : code **CHARLES BABBAGE** ; décode **WKMRSXO WOMKXSAEO**.

(b) Pour un décalage de 13 : code **ZEROS ET UNS** ; décode **TRBETR OBBYR**. Quelle est la particularité du codage et du décodage lorsque le décalage vaut 13 ?

Activité 2 (Attaque du code de César)**.**
On se place dans la peau d'un espion qui vient d'intercepter le message :

NCAN XD WN YJB NCAN

mais qui ne connaît pas le décalage qui a servi à le coder.
Il existe seulement 26 décalages possibles : le code de César n'est pas très sécurisé... C'est un peu long à tester pour un humain, mais très facile pour un ordinateur.
Il existe une autre façon d'attaquer le message codé, car une même lettre est toujours codée de la même façon. Par exemple, pour un décalage de 3, la lettre **A** devient toujours **D** et **E** devient toujours **H**. Or, dans un long texte, les lettres n'apparaissent pas toutes avec la même fréquence.
Pour la langue française, les lettres les plus rencontrées sont dans l'ordre :

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

avec les fréquences :

E	S	A	I	N	T
14,5%	8%	7,5%	7%	7%	7%

Si la lettre **E** apparaît très souvent dans le message de départ, alors pour un décalage de 3, la lettre **H** apparaîtra très souvent dans le message codé.
Voici donc la méthode pour décrypter un message :

- On cherche la lettre la plus fréquente dans le message codé (imaginons que c'est le **K**).
- On suppose que cette lettre correspond au **E**.
- On calcule le décalage (pour aller de **E** à **K** c'est 6).
- On essaie de décoder le message sur la base de ce décalage.
- Si on obtient un message incohérent, on recommence en supposant que la lettre la plus fréquente correspond à l'une des lettres suivantes **S**, **A**, **I**, **N**, **T**.

Avec cette méthode, décrypte les messages suivants (chaque message a été codé avec un décalage différent) :

1. **NCAN XD WN YJB NCAN**
2. **DFC WPD PALFWPD OPD RPLYED**
3. **QSMRW KVERH IX TPYW MRXIPPMKIRX**

Activité 3 (Codage des caractères)**.**
Les ordinateurs préfèrent les nombres aux lettres ! Chaque caractère est numéroté. Voici la table ASCII des premiers caractères.

33	!	43	+	53	5	63	?	73	I	83	S	93	]	103	g	113	q	123	{
34	"	44	,	54	6	64	@	74	J	84	T	94	^	104	h	114	r	124	|
35	#	45	-	55	7	65	A	75	K	85	U	95	_	105	i	115	s	125	}
36	$	46	.	56	8	66	B	76	L	86	V	96	‘	106	j	116	t	126	~
37	%	47	/	57	9	67	C	77	M	87	W	97	a	107	k	117	u	127	-
38	&	48	0	58	:	68	D	78	N	88	X	98	b	108	l	118	v		
39	’	49	1	59	;	69	E	79	O	89	Y	99	c	109	m	119	w		
40	(	50	2	60	<	70	F	80	P	90	Z	100	d	110	n	120	x		
41	)	51	3	61	=	71	G	81	Q	91	[	101	e	111	o	121	y		
42	*	52	4	62	>	72	H	82	R	92	\	102	f	112	p	122	z		

Par exemple, le caractère numéro 37 est le symbole « % ». Le caractère 65 est la lettre majuscule « A ». Le caractère 107 est la lettre minuscule « k ». Les caractères associés aux entiers de 0 à 32 ne sont pas des caractères imprimables.

1. Quelles phrases sont codées par les nombres suivants ?
   (a) 66-111-110-106-111-117-114-33
   (b) 50-43-51-61-53
   (c) 83-101-114-118-105-114 101-110 115-105-108-101-110-99-101-46 (La devise de la NSA.)
2. Trouve l’équivalent numérique des caractères des noms suivants : Boole, Godel, Turing.
3. - On note `chr` la fonction qui à un nombre associe le caractère correspondant. Par exemple `chr(65)` renvoie le caractère « A ».
   - On note `ord` la fonction qui à un caractère associe son numéro dans la table ci-dessus. Par exemple `ord('a')` renvoie l’entier 97.

(a) Que donnent les instructions suivantes : `chr(100)`, `ord('H')`, `chr(65+10)`, `ord(chr(77))`, `chr(ord('#'))` ?

(b) Que fait la suite d'instructions suivantes ? (Essaye d'abord avec la lettre « a ».)

- Entrée : un caractère, noté `car`, parmi « a », « b »,..., « z »
- `n` ← `ord(car)`
- `n` ← `n - 32`
- Sortie : `chr(n)`

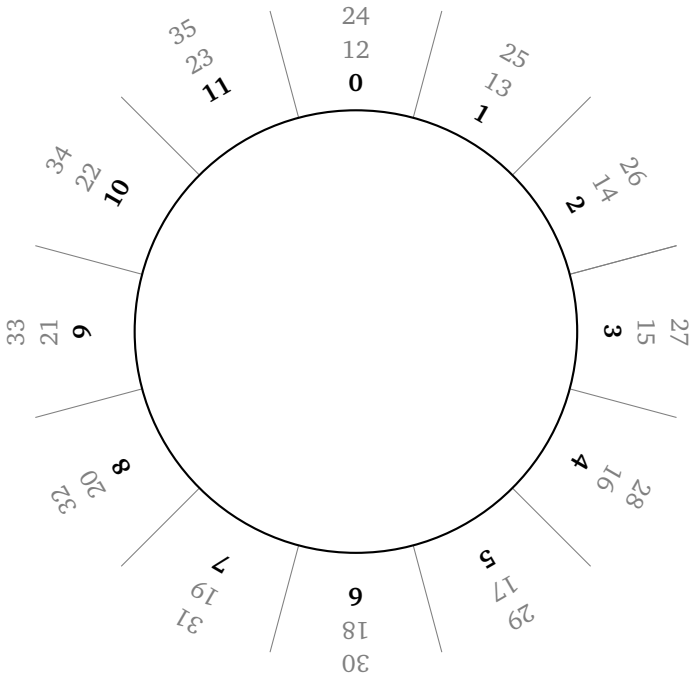
Activité 4 (Modulo)**.**
Compter modulo n, c'est compter uniquement avec les entiers 0, 1, 2,..., $n-1$.

- **Exemple modulo** 60**.** Compter modulo 60, c'est compter comme les minutes d'une montre : 0, 1, 2,..., 59. Quand on arrive à 60, on repart immédiatement à 0. On va noter :

$$60 \pmod{60} = 0 \qquad 61 \pmod{60} = 1 \qquad 62 \pmod{60} = 2 \qquad \ldots$$

- **Exemple modulo** 12**.** Si on compte modulo 12, alors on se ramène à un entier parmi 0, 1,..., 11. On peut s'aider d'une roue pour visualiser :

$$16 \pmod{12} = 4 \qquad 29 \pmod{12} = 5 \qquad 34 \pmod{12} = 10$$

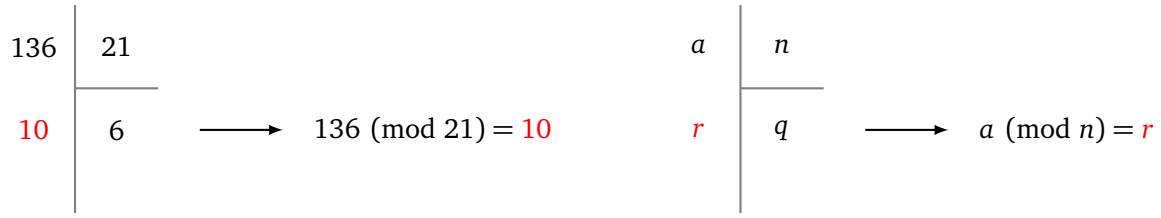


1. (a) Calcule $21 \pmod{12}$; $32 \pmod{12}$; $50 \pmod{12}$; $100 \pmod{12}$.

   (b) Calcule $75 \pmod{60}$; $128 \pmod{60}$; $666 \pmod{60}$.

   (c) Calcule $32 \pmod{26}$; $42 \pmod{26}$; $111 \pmod{26}$.

2. On calcule $a \pmod{n}$ comme le reste de la division euclidienne de a par n. Par exemple pour calculer $136 \pmod{21}$, on écrit la division euclidienne de 136 par 21 : $136 = 6 \times 21 + 10$. Donc $136 \pmod{21} = 10$.



(a) Calcule $1\,254 \pmod{12}$.

(b) Calcule $5\,678 \pmod{60}$.

(c) Calcule $32\,158 \pmod{26}$.

(a) **Modulo 2.**

Calcule $3 \pmod 2$; $4 \pmod 2$; $5 \pmod 2$… Complète et retiens les énoncés suivants :

$a \pmod 2 = 0$ lorsque a est _____

$a \pmod 2 = 1$ lorsque a est _____

(b) **Modulo 10.**

Calcule $21 \pmod{10}$; $39 \pmod{10}$; $2345 \pmod{10}$. Complète et retiens l'énoncé suivant :

$a \pmod{10}$ est le _____ de l'entier a.

(c) **Modulo n.**

Calcule $(12 \times 7) \pmod 7$; $66 \pmod{11}$; $72 \pmod 9$. Complète et retiens :

n divise a exactement lorsque _____

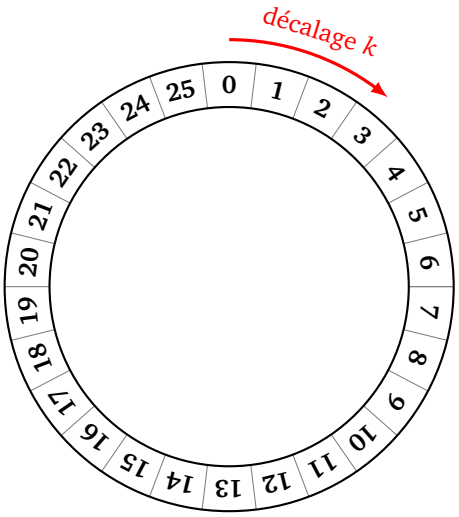
3. **Retour au code de César.**

Le code de César est en fait une simple addition ! Numérotons les lettres par leur rang : « A » est de rang 0, « B » est de rang 1,…, « Z » est de rang 25. Le code de César de décalage 3 consiste simplement à ajouter 3 au rang de la lettre. Comme le rang de la lettre ne peut atteindre 26, il faut calculer le résultat modulo 26. La formule pour un décalage de 3 est donc :

$$\text{rang codé} = \text{rang} + 3 \pmod{26}$$

Pour le code de César de décalage k, la formule est :

$$\text{rang codé} = \text{rang} + k \pmod{26}$$



(a) Calcule le rang codé par un décalage de 8 pour les lettres de rang 7, 15, 23.

(b) Complète les lignes du tableau ci-dessous en suivant le modèle de la première ligne : la lettre « C » est de rang 2 ; avec un décalage de 3, le rang codé est $2+3 \pmod{26} = 5$; la lettre codée est donc le « F ». Attention ! le plus petit rang est 0 (lettre « A »).

Lettre	Rang	Décalage k	Rang codé	Lettre codée
C	2	3	5	F
F		3		
Y		3		
H		15		
	10			R
T			2	
		10		E

4. **Algorithme du code de César.**

Écris un algorithme qui, pour un décalage k fixé, prend en entrée une lettre majuscule et renvoie en sortie la lettre codée par le code de César de décalage k. Par exemple si $k = 3$ et si la lettre entrée est « A » alors la sortie doit être « D ». Les étapes de l'algorithme sont :

- prends le caractère en entrée et transforme-le en un entier compris entre 65 et 90 (voir l'activité précédente) ;
- transforme cet entier en un nombre entre 0 et 25 ;
- applique la formule du décalage de César sur cet entier ;
- revient à un entier entre 65 et 90, puis à un caractère.

Activité 5 (Le chiffrement de Vigenère)**.**

Le code de César n'est pas assez sûr, le chiffrement de Vigenère en est une version plus sophistiquée. Par exemple pour coder :

ALLEZ ASTERIX

- on commence par découper notre texte en blocs de même longueur, par exemple de longueur 3 :

ALL EZA STE RIX

- on choisit une clé, composée de 3 nombres, par exemple $(3, 6, 5)$;
- on décale de 3 la première lettre de chaque bloc (le **A** du premier bloc devient **D**) ;
- on décale de 6 la deuxième lettre de chaque bloc (le premier **L** du premier bloc devient **R**) ;
- on décale de 5 la troisième lettre de chaque bloc (le second **L** du premier bloc devient **Q**) ;
- on obtient par blocs **DRQ HFF VZJ UOC** et le message codé est alors :

DRQHF FVZJUOC

Note les deux améliorations par rapport au code de César classique :

- une même lettre peut être codée de plusieurs façons différentes (par exemple le premier **L** est codé en **R** alors que le second est codé en **Q**) ;
- une même lettre du message codé peut correspondre à différentes lettres du message initial (par exemple le premier **F** code un **Z** alors que le second **F** code un **A**).

1. Code la phrase **RIEN NE SE PERD** avec la clé $(3, 6, 5)$.
2. Décode la phrase **WUZW YJ WXFQYKRXRH** codée avec la clé $(3, 6, 5)$.
3. Code la phrase **EQUATEUR** avec la clé $(1, 25, 10, 5)$.
4. Décode la phrase **NDBNEHOS**, codée avec la clé $(1, 25, 10, 5)$.
5. Décode la phrase suivante d'Albert Einstein, par une attaque par fréquence, sachant qu'elle a été codée par un chiffrement de Vigenère avec une clé de longueur 3 :

NE CKI J GWA ESTOI BPI IKGFEPLVXL
KP MCYA CZHPGLT TVWV UG THU TLTHYG P LSYPNMITI