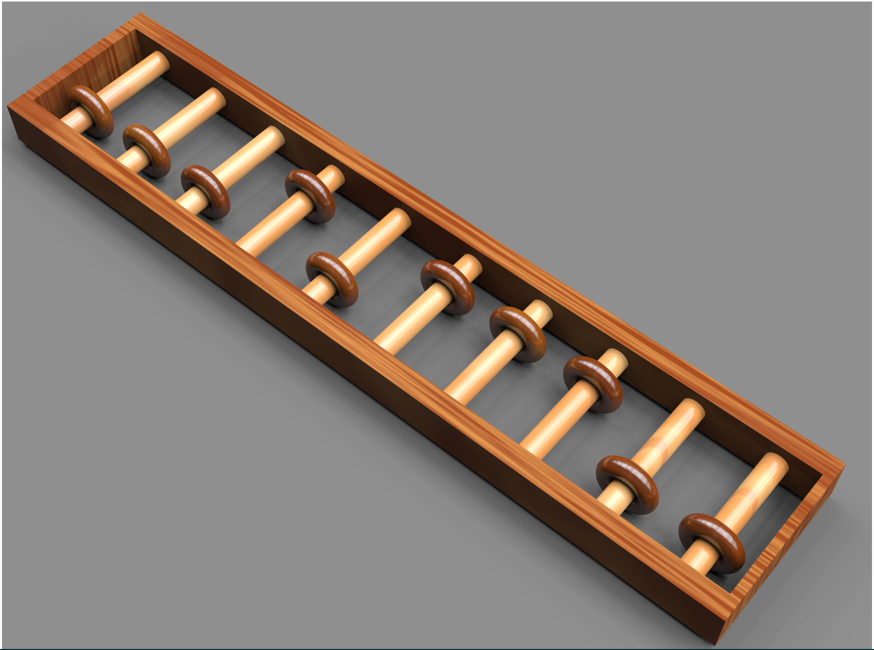
Codage de l'information

Laurent Noé-Léopold Weinberg
paternité : Francesco De Comité

Licence 1 Mathématiques Informatique Semestre 1
Université de Lille
Faculté des Sciences et Technologies

10 septembre 2025



But du cours

Objectifs

- Découvrir les informations élémentaires manipulées par les ordinateurs.
- Panorama des différentes façons de transmettre des informations de base.
- Applications à l'informatique : comment les ordinateurs stockent, traitent et transmettent les informations de base.

Déroulement du cours

Plan

- Qu'est-ce que l'information ?
- Quelles informations ?
  - Les nombres (naturels, relatifs, flottants)
    - Histoire (succinte)
    - Codage
    - Operations et manipulations.
  - Les caractères et les textes.
  - Autres types d'informations (images, sons, fichiers...)

Communiquer

Communiquer

- La vie en communauté implique que ses membres puissent communiquer entre eux.
- Partager des informations :
  - Un évenement est intervenu.
  - J'ai besoin d'aide.
  - Je propose un échange.
- Pour communiquer, il faut un langage commun, une façon commune de représenter les choses.
- Différents domaines de communication $\Rightarrow$ différents *langages* $\Rightarrow$ différents *codages*.

Les codages élémentaires

Informations basiques

- Parler, écrire : coder des textes (des caractères ? $\rightarrow$ pas toujours !).
- Recenser, commercer, compter : coder des quantités (des chiffres et des nombres).
- Données visuelles : des plans, des schémas : coder des images.

Codages vus dans ce cours

Limitations

- On se limitera dans ce cours à deux codages principaux :
  - Les nombres.
  - Les caractères.
- Pourquoi ?
  - Les ordinateurs manipulent principalement (et historiquement) des nombres.
  - Les différentes architectures d'ordinateurs privilégient la manipulation rapide et efficace des nombres.
  - Le codage des caractères pourra être vu comme un exemple de codage et de transmission de données hétéroclites.
  - Ce codage est alors un modèle pour le codage d'autres informations (sons, images, données diverses...)

Coder les nombres

Supputations historiques

Utilisations primitives des nombres

- Compter et mémoriser :
  - Taille de la famille (pas forcément besoin de nombres)
  - Taille du troupeau : savoir ce que l'on possède.
- Définir les bases d'un échange :
  - Sur la base " un contre un" : pareil que compter.
  - Lorsque chaque objet de l'échange a une *valeur* différente.
- Calculer :
  - Définir le montant de l'impôt (surface d'un champ ...)
  - Calculer et comparer des volumes.
  - Partager un butin, la production d'une activité (agricole, artisanale, industrielle)

Supputations historiques (suite et fin)

Utilisations moins primitives des nombres

- Développement des mathématiques : résoudre des problèmes arithmétiques / numériques.
- Acquérir, représenter et transmettre des résultats de mesure (sciences experimentales).
  - Mise au point d'une norme pour chaque mesure.
  - Eviter les erreurs de transmissions : codes sans erreurs, ou capables de corriger leurs erreurs.
  - Minimiser la bande passante : codes compacts.
  - Tsunami de données : capacités de traitement et de stockage, automatisation du traitement.

Représenter des quantités

Le bâton d'Ishango

crédit : Muséum d'Histoire Naturelle Bruxelles

Le bâton d'Ishango

Description

- Age : 20.000 ans, découvert en 1950 au Congo.
- Quatre groupes d'encoches, manifestement des représentations de nombres.
- Et puis c'est tout : les interprétations les plus diverses existent ...

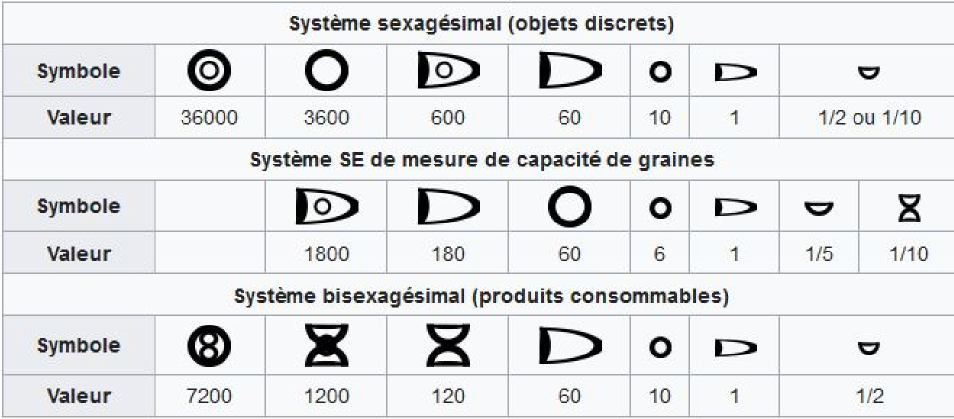
Remarques

- La plus vieille représentation de nombres connue.
- Un nombre est représenté par un groupe de lignes verticales (des *batônnets*)

Remarques

- Troisième millénaire avant J.C.
- Système utilisé dans les transactions commerciales (plus ou moins équivalents à des bons de livraisons).
- Plusieurs systèmes selon le type d'articles à dénombrer.
- Les nombres se lisent en additionnant les symboles (numération additive).

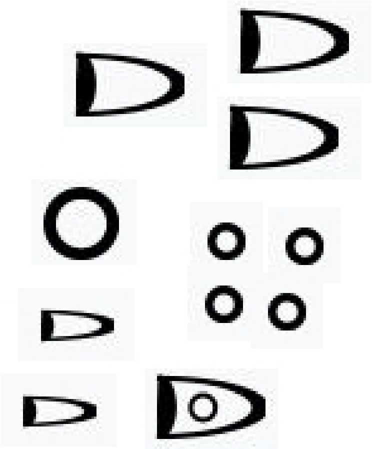
Sumer et Babylone



Système sexagésimal (objets discrets)								
Symbole								
Valeur	36000	3600	600	60	10	1	1/2 ou 1/10	
Système SE de mesure de capacité de graines								
Symbole								
Valeur		1800	180	60	6	1	1/5	1/10
Système bisexagésimal (produits consommables)								
Symbole								
Valeur	7200	1200	120	60	10	1	1/2	

crédit : Wikipedia

Sumer : exemple de nombre



$$3600 + 600 + (3 * 60) + (4 * 10) + (2 * 1) = 4422$$

Sumer : exemple de nombre

Remarques

- Pas de positions définies pour les *chiffres*
- Pas adapté aux calculs tels que nous les pratiquons (ni à d'autres façons de calculer...).

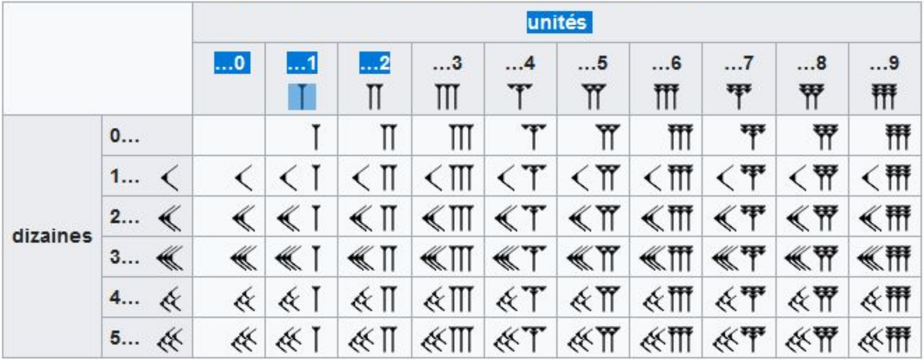
Babylone

crédit : Wikipedia

Chiffres cunéiformes (*étymologie : en forme de coin*)

- Cinquante-neuf chiffres différents.
- Un motif qui se répète pour chaque dizaine.
- Pas de symbole pour zéro.

Liste des chiffres cunéiformes babyloniens de 0 à 59.



crédit : Wikipedia

Numération cunéiforme : quelques exemples

Exemples de nombres écrits en numération babylonienne sexagésimale.

Valeur décimale	Écriture babylonienne cunéiforme	Décomposition en base 60
1		1 x 1
17		17 x 1
44		44 x 1
60		60 = 1 x 60 + 0 x 1
85		1 × 60 + 25 x 1
3600		3600 = 1 x 60² + 0 x 60 + 0 x 1
11327		3 × 60² + 8 × 60 + 47 x 1
7000,2525		1 x 60² + 56 x 60 + 40 x 1 + 15/60 + 9/60²

crédit : Wikipedia

- On note que, par exemple, 1 et 3600 s'écrivent de la même façon.
- L'ambiguité disparaît quand tous les chiffres sont présents.

Rome antique

Les chiffres

1	5	10	50	100	500	1000
I	V	X	L	C	D	M

Exemples

1	4	6	9	12	40	80	100
I	IV	VI	IX	XII	XL	LXXX	C

Remarques

- Numération quasi-additive :
  - Des symboles particuliers pour chaque dizaine.
  - Ordre dans les puissances de 10.
- Utile pour dénombrer et indiquer des dates, impraticable pour les calculs.

Rome antique : calculer

crédit : Goscinny/Uderzo

Rome antique : calculer ?

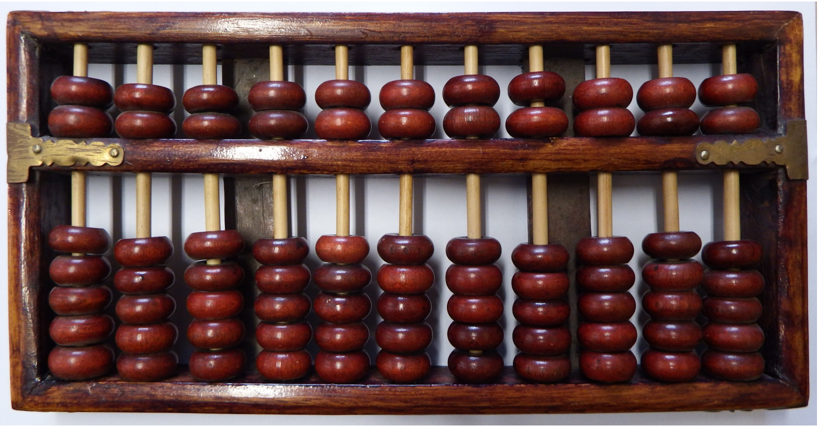
Méthodes

- Pour calculer, les Romains utilisaient des creux dans le sable et des cailloux.
- Table de poussière : *abacus*.
- petit caillou : *calculus*.
- Les tables de poussière et les cailloux sont remplacés par un instrument : l'*abaque*.

Rome antique : calculer

crédit : Joern Luetjens

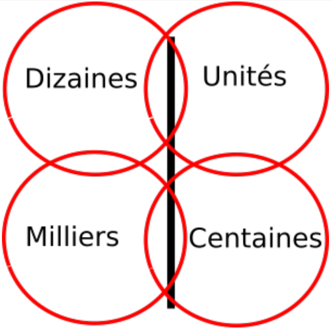
Ailleurs : boulier chinois



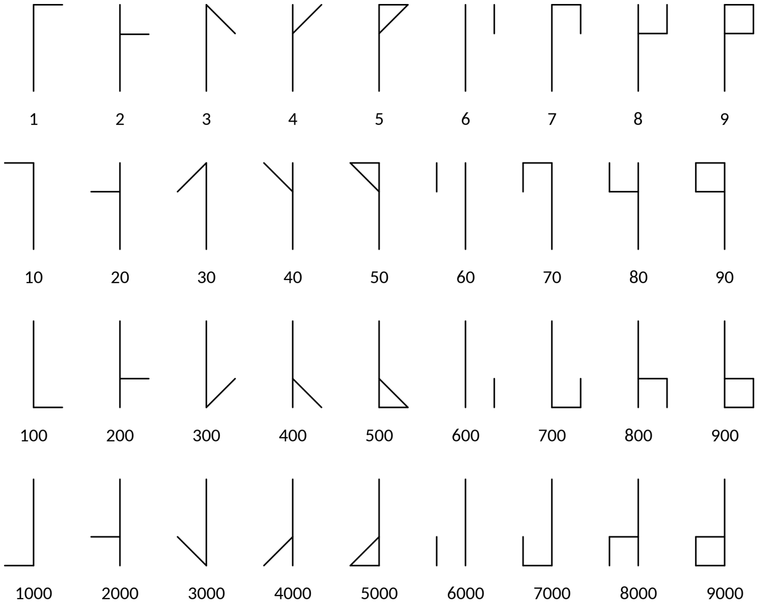
Un dernier pour la route : Notae Elegantissimae

Description

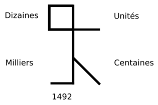
- Système de notation numérique (donc pas un système de calcul), utilisé par les moines cisterciens au XIII ème siècle.
- Bâti autour d'une barre verticale, divisant l'espace en quatre quadrants.
- Chaque cadran correspond à une puissance de 10.

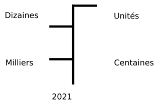


Un dernier pour la route : Notae Elegantissimae



Notae Elegantissimae : exemples





La représentation moderne des nombres

Caractéristiques

- Apparaît en Europe à la Renaissance.
- Numération de position : la place du chiffre dans le nombre donne sa valeur.
- Utilisation du zéro.
- Dix chiffres : base 10
- Permet le développement du calcul mathématique.

La numération de position en base 10

Définition

- On définit un ensemble $\mathcal{C}$ de 10 caractères, les chiffres.
- Chacun de ces chiffres est associé à une valeur unique comprise entre 0 et 9.
- Un nombre est représenté par une suite de chiffres

$$a_{t-1} \cdots a_1 a_0, \forall i \in [\![0, t-1]\!], a_i \in \mathcal{C}$$

- La valeur représentée par ce nombre est :

$$n = \sum_{i=0}^{t-1} a_i \times 10^i$$

La numération de position en base 10

Remarques

- Très proche de la numération cunéiforme.
- Apparition (invention) du zéro.
- Se prête bien aux algorithmes des opérations de base (addition, multiplication) ...à moins que ce ne soit l'inverse : nous sommes tellement habitués à appliquer ces algorithmes que nous pensons qu'ils sont simples.

Notations et définitions

- La taille d'un nombre est définie par le nombre de chiffres qui le composent.
- Le chiffre représentant les unités, le plus à droite dans l'écriture du nombre est appelé le *chiffre de poids faible*.
- Réciproquement, le chiffre le plus à gauche est appelé le *chiffre de poids fort*.
- Le chiffre de poids fort est celui qui a le plus d'influence sur la valeur finale du nombre.
- Le chiffre de poids fort est forcément différent de zéro.
- On dira que des zéros éventuels à gauche du chiffre de poids fort sont *non significatifs*.

La numération de position en base 10

Propriétés

- La décomposition d'un nombre en base 10 est unique. (cf exercices).
- On peut comparer facilement deux nombres de taille quelconque (cf exercices).
- Quel est le nombre de chiffres de l'écriture d'un nombre en base 10 ?

Calculer

Quelles opérations ?

Opérations basiques

- A priori les quatre opérations :
  - Addition
  - Soustraction
  - Multiplication
  - Division
- Ce sont les plus courantes.
- On a développé des algorithmes et des moyens mnémotechniques largement diffusés, et enseignés dès le plus jeune âge.
- Les autres opérations (extraction de racine carrée, trigonométrie, exponentielles) sont moins courantes, on peut les réserver à des spécialistes, ou bien écrire des algorithmes ou bien utiliser des bibliothèques lorsqu'on en aura besoin.

Quelles opérations ?

Restriction du domaine

- En y regardant de plus près, on peut se concentrer sur deux opérations : l'addition et la multiplication.
- Parce que leurs algorithmes sont plus simples, plus rapides ,et peuvent être facilement implémentés dans les ordinateurs, soit en les câblant, soit en les programmant.
- Parce que la soustraction pourra se ramener à une addition, avec un algorithme plus simple que celui appris à l'école primaire.
- Parce la division est un algorithme beaucoup plus compliqué (voir la complexité de son apprentissage à l'école primaire).

Différentes solutions ?

Selon le système de numération ?

- La notation sumérienne, les chiffres romains *nous semblent* mal adaptés à la pratique de l'addition (essayez...).
- Pour les nombres en notation cunéiforme, on a un problème dû à l'absence du zéro (compréhension du nombre et positionnement des chiffres)
- Conclusion : à première vue, la notation positionnelle semble la mieux adaptée à l'exercice de l'addition.
- Mais attention au biais : on sait peu de choses sur la pratique des calculs élémentaires dans les autres civilisations.

Comment ça marche ?

L'algorithme de l'addition

1. Soit à additionner les deux nombres $A = \overline{\mathtt{a_n a_{n-1} \cdots a_1 a_0}}$ et $B = \overline{\mathtt{b_m b_{m-1} \cdots b_1 b_0}}$
2. Superposer les deux nombres, en alignant leur chiffre de poids faible (*les unités*).
3. Si les deux nombres sont de longueur différentes, compléter (mentalement ou pratiquement) le nombre le plus court par des zéros à gauche, jusqu'à obtenir deux nombres de même longueur.
4. En commençant par les chiffres des unités, *additionner les deux chiffres*, poser le chiffre des unités du résultat, et reporter la retenue éventuelle sur la colonne suivante à gauche.

Comment ça marche ? Un exemple

	1	5	4	6
+			5	7

	1	5	4	6
+	0	0	5	7

	1	5	4	6
+	0	0	5	7
				13

			1	
	1	5	4	6
+	0	0	5	7
				3

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

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

Analyse (succinte) de l'algorithme

Analyse

- On n'a besoin de connaître *par cœur* que les additions de deux chiffres : les tables d'addition.
- La retenue vaut au plus 1.
- L'algorithme est une suite d'additions élémentaires (deux chiffres et une éventuelle retenue).
- La taille du résultat est soit T, soit $T + 1$. Dans ce deuxième cas, le chiffre de poids fort vaut 1.
- On est très proche de l'algorithme vu en Architecture Elémentaire.
- Complexité en temps de l'addition ?

Complexité en temps de l'addition

Analyse

1. La première étape est de ramener les deux nombres à la même longueur, en ajoutant des zéros à gauche du nombre ayant le moins de chiffres. Soit T la taille des nombres après cette étape.
2. On effectue ensuite T fois une addition élémentaire.
3. On pourrait arrêter le calcul lorsque l'on a dépassé le dernier chiffre significatif du nombre le plus court, et qu'il n'y a plus de retenue à propager, mais on gagne peu de chose.
4. La complexité en temps est donc proportionnelle à la longueur du nombre le plus long : complexité linéaire en T.
5. Il est impossible d'améliorer significativement cette complexité, puisqu'il faudra de toute façon lire chacun des chiffres de l'addition au moins une fois.

- On pose les chiffres comme pour l'addition.
- Pas besoin d'ajouter des chiffres non significatifs.
- 207 est le *multiplicateur*.
- 3546 est le *multiplicande*.

	3	5	4	6
$\times$		2	0	7

- On commence par multiplier le multiplicande par le chiffre de poids faible du multiplicateur.
- Pour cela, on multiplie chaque chiffre du multiplicande, de la droite vers la gauche, par le chiffre de poids faible du multiplicateur.
- On pose le chiffre des unités de cette multiplication de deux chiffres, et on reporte une éventuelle retenue vers la colonne suivante.

			4	
	3	5	4	6
$\times$		2	0	7
				2

La multiplication

- On continue pour tous les chiffres du multiplicande.
- On se prépare à recommencer pour le chiffre suivant du multiplicateur, mais le résultat deette multiplication élémentaire sera décalé d'un rang vers la gauche (correspondant à une multiplication par 10), symbolisé par l'écriture d'un point.
- Ici, la deuxième multiplication élémentaire a pour résultat zéro.

	3	5	4	6
$\times$		2	0	7
2	4	8	2	2
0	0	0	0	·

La multiplication

- On recommence avec le troisième chiffre du multiplicateur, en décalant deux fois le résultat.

		3	5	4	6
	×		2	0	7
	2	4	8	2	2
	0	0	0	0	·
7	0	9	2	·	·

La multiplication

- On additionne les résultats intermédiaires, ce qui donne le résultat final.

		3	5	4	6
	$\times$		2	0	7
	2	4	8	2	2
	0	0	0	0	·
7	0	9	2	·	·
7	3	4	0	2	2

Analyse de la multiplication

Remarques

- Lors d'une multiplication élémentaire, la retenue engendrée est inférieure ou égale à 8.
- Le résultat de la multiplication du multiplicande par un chiffre a une taille égale ou supérieure de 1 à celle du multiplicande.
- Lorsque le chiffre du multiplicateur vaut 0, on peut remplacer la ligne par un décalage supplémentaire.
- L'addition finale peut être remplacée par une suite d'additions intermédiaires, sans modifications notables de l'algorithme.

Complexité de la multiplication

Complexité

- Le coût de la multiplication du multiplicande par un chiffre est proportionnel au nombre de chiffres du multiplicande (notons-le M).
- Cette multiplication est répétée autant de fois qu'il y a de chiffres au multiplicateur (notons-le N).
- Le tout est suivi d'une addition, dont la complexité est de l'ordre de M, répété N fois.
- La complexité de cet algorithme est donc de l'ordre de $M \times N$
- Si les deux nombres ont la même longueur, on a une complexité en N^2, donc quadratique : si la taille des nombres double, l'algorithme prend quatre fois plus de temps.

Complexité de la multiplication

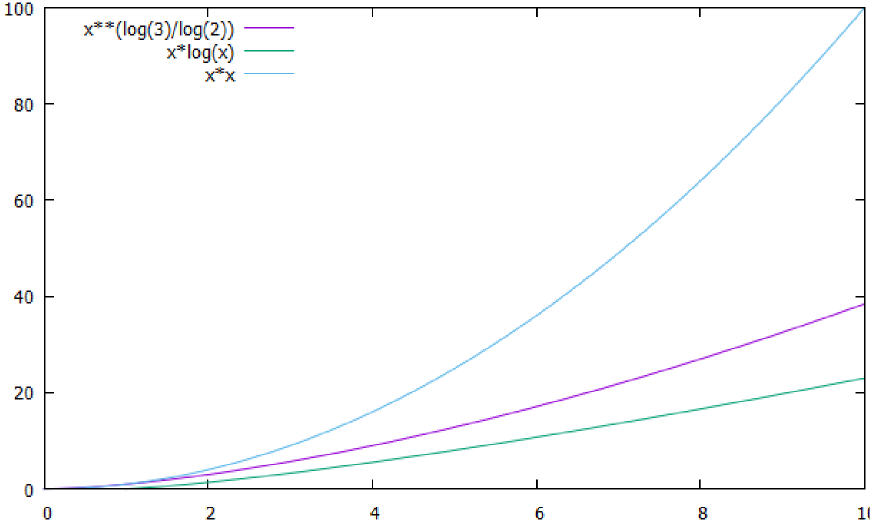
Pour aller plus loin ...

- La multiplication à l'aide d'un boulier suit le même algorithme, avec des contraintes de place : il faut bien prévoir la place des résultats intermédiaires.
- Il existe des algorithmes plus efficaces (c'est-à-dire meilleurs que quadratiques en temps) :
  - L'algorithme de Karatsuba (1962) a une complexité en

    $$\mathcal{O}(N^{\log_2 3}) \approx \mathcal{O}(N^{1.58})$$

    Si la taille du nombre double, le temps de calcul est multiplié par 3 ($2^{1.58}$).
  - En mars 2019, Harvey et Van der Hoeven décrivent une méthode en $\mathcal{O}(n \log n)$. On *conjecture* qu'il n'y a pas de méthode plus rapide.
  - Conclusion : même pour les algorithmes soi-disant élémentaires, la recherche est active.

Comparaison des complexités



La soustraction

Algorithme

- On apprend un algorithme de soustraction à l'école primaire.
- Il existe un autre algorithme :
  - Simple à définir et à utiliser.
  - Qui se transpose facilement dans des circuits d'ordinateur.
  - Qui utilise l'algorithme d'addition : du point de vue technique, on pourra réutiliser un circuit déjà existant.
- Cet algorithme se base sur la représentation en *complément à 9* des entiers négatifs.

Complément à 9 (à 10)

Comment ça marche ?

- Soit à effectuer la soustraction $845 - 167 = N$
- Les deux nombres s'écrivant sur trois chiffres, le résultat N aura au plus trois chiffres.
- On peut écrire $1000 + 845 - 167 = 1000 + N$
- En réarrangeant les termes : $845 + (1000 - 167) = 1000 + N$
- Ou encore : $845 + (999 - 167 + 1) = 1000 + N$
- $999 - 167 + 1$ se calcule facilement en remplaçant chacun des chiffres de 167 par leur *complément à neuf* : 832, puis en ajoutant 1 ($\longrightarrow$ 833)
- La soustraction de départ devient une addition : $845 + 833 = 1678$
- On retranche le 1000 qu'on avait ajouté de chaque côté de l'égalité, ce qui donne le résultat 678.

Soustraction : analyse de la méthode

Remarques

- Il faut que les deux nombres soient de la même longueur.
- Et si ça n'est pas le cas ? On complète le nombre le plus court par des zéros à gauche.
- Et si le deuxième nombre est plus grand que le premier (résultat négatif) ?

Soustraction mode d'emploi

OSolution

- On s'arrange pour soustraire le nombre le plus petit du plus grand.
- On fixe le signe du résultat en fonction de la magnitude des deux opérandes.

Exemple

- Soit la soustraction $247 - 1023$
- On sait (et on note quelque part) que le résultat sera négatif.
- On calcule $1023 - 247$ par la méthode du complément à 10 :
  - On remplace -247 par $1000 - 247 = 753$
  - On effectue l'addition $1023 + 753 = 1776$
  - On soustrait le 1000 ajouté précédemment : $1776 - 1000 = 776$
  - On change le signe du résultat : $247 - 1023 = -776$

Calculer : conclusion

Résumé

- On a décrit les algorithmes élémentaires des opérations de base : l'addition et la multiplication.
- Les versions étudiées de ces algorithmes sont bien adaptées à la notation positionnelle en base 10.
- La soustraction a elle aussi un algorithme spécifique, mais on peut s'en passer et utiliser l'algorithme de l'addition.
- C'est avec ces algorithmes en tête qu'on abordera la partie suivante du cours : le traitement des nombres par les ordinateurs.

Calculer : conclusion

Aspects non étudiés jusqu'ici

- La division reste un algorithme plus difficile, on l'ignore, au moins pour le moment.
- D'autres algorithmes existent dans d'autres cultures et d'autres civilisations, on ne les étudiera pas.