

# Chapitre 1 - Ensembles de mots

Benjamin WACK (cours) - Mica MURPHY (note) - Antoine SAGET (note)

Lundi 1er Octobre 2018

## 0) Introduction

Discret est l'opposé de continu, et il peut y avoir un nombre fini ou infini de valeurs. On ne fera ni de géométrie ni d'analyse de fonctions (dérivées, etc.).

## 1) Mots

### a) Alphabets et mots

**Définition.** Un **alphabet** est un ensemble fini de symboles.

*Exemples.*

- alphabet de 26 lettres
- code ASCII
- notes de musique

**Définition.** un **mot** sur un alphabet  $A$  est une suite ordonnée finie de symboles de  $A$ .

L'ordre des lettres est important : abba est différent de baab. Il peut y avoir des répétitions.

Si  $x_1, x_2, x_n$  sont des symboles de  $A$  ; on peut parler du mot  $x = x_1x_2...x_n$

**Cas particulier.** Le **mot vide** à 0 symboles noté  $\epsilon$ .

$\epsilon$  n'est pas un symbole de  $A$

On note  $A^n$  l'ensemble des mots sur  $A$  formés de  $n$  symboles et  $A^*$  l'ensemble de tous les mots sur  $A$ .

**Définition.** On appelle **longueur d'un mot** le nombre de symboles qui le composent.

$$lg(x_1x_2...x_n) = n$$

$$lg(\epsilon) = 0$$

Dans  $A^*$  on retrouve chaque symbole de  $A$  sous la forme d'un mot de longueur 1.

*Exemples.*

- alphabet latin à 26 lettres  
Toute suite de lettres est appelée **mot** (même s'il n'est pas dans le dictionnaire)
- alphabet binaire  $B = \{0, 1\}$  Il y a  $2^n$  mots binaires de longueur  $n$ .
- alphabet des chiffres  $\{0, 1, 2, \dots, 9\}$  un mot sur cet alphabet représente un nombre entier

**Définitions.** On appelle **langage** sur  $A$  un ensemble (fini ou infini) de mots sur  $A$ , autrement dit une partie de  $A^*$ .

*Exemples.*

- Les mots du dictionnaire *Larousse 2018*
- Les suites de chiffres qui ne commencent pas par un 0.
- Le langage d'un seul mot  $\{u\}$
- $\{\epsilon\}$
- Le langage vide :  $\{\emptyset\}$  (à ne pas confondre avec  $\epsilon$  !)
- $A^*$

## b) Préfixe, suffixe, facteur

### Concaténation

Soient  $u = u_1u_2 \dots u_n$  et  $v = v_1v_2 \dots v_p$  alors le **concaténé** de  $u$  et  $v$  noté simplement  $uv$  est le mot  $u_1u_2 \dots u_nv_1v_2 \dots v_p$

*Exemple.* Si  $u = 1011$  et  $v = 010$  alors  $uv = 1011010$

### Préfixe, suffixe, facteur

Soient  $u$  et  $v$  deux mots sur  $A$ . On dit que  $u$  **est un préfixe de**  $v$  si il existe un mot  $w$  tel que  $v = uw$   
 $w$  peut être le mot vide.

On note  $u \sqsubseteq v$  le fait que  $u$  est préfixe de  $v$   $u \sqsubset v$  le fait que  $u$  est préfixe strict de  $v$  (cas où  $w \neq \epsilon$ )

**Autre caractérisation :** si  $u = u_1u_2 \dots u_n$ ,  $v = v_1v_2 \dots v_p$  alors  $u \sqsubseteq v$  si et seulement si  $u_1 = v_1, u_2 = v_2, \dots, u_n = v_n$  et  $n \leq p$

**Propriété.** Si  $u \sqsubseteq v$  et  $v \sqsubseteq u$  alors  $u = v$

**Propriété.** Si  $u \sqsubseteq v$  alors  $lg\ u \leq lg\ v$  et si  $u \sqsubset v$  alors  $lg\ u < lg\ v$

On dit que  $u$  est un :

- **suffixe** de  $v$  s'il existe un mot  $w$  tel que  $v = wu$ .
- **facteur** de  $v$  si il existe 2 mots  $x$  et  $y$  tels que  $v = xuy$

*Exemples.* Soit le mot  $baaca$  :

- ses préfixes sont  $\epsilon, b, ba, baa, baac, baaca$ .
- ses suffixes sont  $\epsilon, a, ca, aca, aaca, baaca$
- ses facteurs sont  $\epsilon, b, ba, baa, baac, baaca, a, aa, aac, aaca, ac, aca, c, ca$

**Propriété.** Si  $u$  est un mot de longueur  $n$ , il admet exactement  $n + 1$  préfixes distincts,  $n + 1$  suffixes distincts et au moins  $n + 1$  facteurs (souvent plus).

**Propriétés.**

- $lg(uv) = lg(u) + lg(v)$
- $lg(u^n) = n \times lg(u)$  (où  $u^n$  est le mot  $u$  répété  $n$  fois)
- $u^0 = \epsilon$

Soit  $P : "w = uv"$  et  $Q : "lg(w) = lg(u) + lg(v)"$  on a  $P \Rightarrow Q$ .

La réciproque ( $Q \Rightarrow P$ ) n'est pas vraie :   
 Si  $w = uv$  alors  $lg(w) = lg(u) + lg(v)$    
 Si  $lg(w) = lg(u) + lg(v)$  alors  $W = uv$    
 Contre-exemple :  $u = a, v = b, w = aa$

En revanche, la contraposée ( $\neg Q \Rightarrow \neg P$ ) est vraie : Si  $lg(w) \neq lg(u) + lg(v)$  alors  $w \neq uv$

### c) Distance entre mots

Soient  $u$  et  $v$  deux mots sur  $A$  de même longueur La **distance** de  $u$  à  $v$  est le nombre de symboles de  $u$  qu'il faut modifier pour obtenir  $v$ .

*Exemples.*

- $u = arbre, v = aller, d(u, v) = 4$  (seul le  $a$  est identique aux 2)
- $u = 0101110, v = 0011101, d(u, v) = 4$  (seuls 3 sur 7 caractères sont identiques aux 2)

**Propriétés.** (qui disent que  $d$  est bien une distance)

- $d(u, v) = 0$  ssi  $u = v$
- $d(u, v) = d(v, u)$
- inégalité triangulaire :  $\forall u, v, w,$

$$d(u, v) \leq d(u, w) + d(w, v)$$

**Preuve.**  $d(u, v) = \sum_{i=1}^n d(u_i, v_i)$ , d'où  $d(u, w) + d(w, v) = \sum_{i=1}^n (d(u_i, w_i) + d(w_i, v_i))$ . On peut donc se focaliser sur un seul symbole à la fois : - si  $u_i = v_i$  alors  $d(u_i, v_i) = 0 \leq d(u_i, w_i) + d(w_i, v_i)$  - si  $u_i \neq v_i$  alors  $d(u_i, v_i) = 1$  et  $w_i$  est différent d'au moins un des deux.  $d(u_i, w_i) + d(w_i, v_i) = 1 + 0$  ou  $0 + 1$  ou  $1 + 1$

## 2) Ordre lexicographique

**Idée :** comme l'ordre du dictionnaire.

Soit  $A$  un alphabet quelconque,  $A = \{a_1, a_2, \dots, a_k\}$

On dit que  $A$  est **ordonné** si on fixe un ordre  $<$  sur les symboles, par exemple,  $a_1 < a_2 < \dots < a_k$

Soient  $u$  et  $v$  deux mots de  $A^*$ ,  $u$  est **avant**  $v$  dans l'ordre lexicographique (noté  $u \leq_{\text{lex}} v$ ) si :

- $u$  est un préfixe de  $v$  OU
- il existe un mot  $w$  et deux symboles  $x < y$  tels que  $wx \sqsubseteq u$  et  $wy \sqsubseteq v$

Autrement dit si  $u = u_1 u_2 \dots u_n$ ,  $v = v_1 v_2 \dots v_p$  :

- $n \leq p$  et  $u_1 = v_1, \dots, u_n = v_n$  OU
- $\exists k$  tel que  $u_1 = v_1, \dots, u_k = v_k$  et  $u_{k+1} < v_{k+1}$

*Remarque.* Si  $u$  et  $v$  sont de longueur 1, les ordres  $\leq$  sur  $A$  et  $\leq_{\text{lex}}$  sur  $A^*$  coïncident.

*Exemple.* Sur  $B = \{0, 1\}$  avec  $0 < 1$ , rangeons tous les mots de longueur  $\leq 3$ : 1.  $\epsilon$  2. 0, 1 3. 00, 01, 10, 11 4. 000, 001, 010, 011, 101, 110, 111

$$\begin{aligned} \epsilon &\leq_{\text{lex}} 0 \leq_{\text{lex}} 00 \leq_{\text{lex}} 000 \leq_{\text{lex}} 001 \leq_{\text{lex}} 01 \leq_{\text{lex}} 010 \leq_{\text{lex}} 011 \\ &\leq_{\text{lex}} 1 \leq_{\text{lex}} 10 \leq_{\text{lex}} 100 \leq_{\text{lex}} 101 \leq_{\text{lex}} 11 \leq_{\text{lex}} 110 \leq_{\text{lex}} 111 \end{aligned}$$

**Propriété.**  $\leq_{\text{lex}}$  est un **ordre total** : quels que soient  $u$  et  $v \in A^*$  on a toujours  $u <_{\text{lex}}$  ou  $u >_{\text{lex}}$  ou  $u = v$ . Par exemple, ce n'est pas le cas pour  $\sqsubseteq$ .

*Remarque.* L'ordre lexicographique n'est pas commode à définir, par contre on peut écrire un **algorithme** pour décider si  $u \leq_{\text{lex}} v$  (cf. TD3)

## 3) Ensembles et dénombrement

### a) Notion d'ensemble, fini ou infini

Un ensemble  $E$  est une collection d'éléments sans ordre ni répétition. Si  $E$  est fini, on peut le décrire explicitement par exemple  $\{a, b, c\}$  ou encore  $\{c, a, b\}$ . Cette notation est limitée : on écrit vite des ensembles comme  $\{a, b, c, \dots\}$ , ce qui est ambigu.

En général on décrit plutôt {la forme générale} de éléments de l'ensemble ou {les propriétés}.

*Exemple.*  $\{2k + 1 | k \in \mathbb{N}\} = \{1, 3, 5, \dots\}$   $\{k \in \mathbb{N} | k \text{ impair}\}$

**Notations.** -  $x \in E$  : l'élément  $x$  **appartient** à l'ensemble  $E$  -  $A \subseteq B$  : l'ensemble  $A$  est contenu / inclus dans  $B$ , autrement dit tout élément de  $A$  appartient à  $B$  -  $A \subset B$  : inclusion stricte (si  $A \subseteq B$  et  $A \neq B$ ) - l'ensemble vide  $\emptyset$  ne contient aucun élément :  $\{\}$

*Exemples.* -  $1 \in \{0, 1\}$  -  $-5 \notin \mathbb{N}$  -  $-5 \in \mathbb{Z}$  -  $\mathbb{N} \subset \mathbb{Z}$  -  $\mathbb{Z} \subseteq \mathbb{Z}$  -  $\{\epsilon\} \neq \emptyset$  -  $\{-1, 1\} \subsetneq \mathbb{N}$

*Remarque.*  $x \in E$  si et seulement si  $\{x\} \subseteq E$

**Définition.** Le **cardinal** d'un ensemble **fini** est le nombre d'éléments qui le composent, noté  $\text{card } E$ ,  $\#E$  ou  $|E|$ .

*Exemples.* -  $\text{card}(\{a, b, \dots, z\}) = 26$  -  $\text{card}(\{0, 1, \dots, 9\}) = 10$  -  $\text{card}(\{\epsilon\}) = 1$  -  $\text{card}(\emptyset) = 0$

**Attention.** Par convention, dans tout ce cours, si on parle du cardinal d'un ensemble, celui-ci est fini

*Propriétés.* - Si  $X \subseteq Y$  alors  $\text{card } X \leq \text{card } Y$  - Si  $X \subset Y$  alors  $\text{card } X < \text{card } Y$

(Notez le parallèle entre  $\subseteq, \leq$  et  $\sqsubseteq$  et entre  $\subset, <$  et  $\sqsubset$ )

## b) Opérations entre ensembles

### Union et intersection

L'**union** de  $X$  et  $Y$  est l'ensemble des éléments présents dans  $X$  ou  $Y$ .

$$X \cup Y = \{x | x \in X \text{ ou } x \in Y\}$$

L'**intersection** de  $X$  et de  $Y$  est l'ensemble des éléments présents la fois dans  $X$  et dans  $Y$ .

$$X \cap Y = \{x | x \in X \text{ et } x \in y\}$$

**Attention :** -  $X$  et  $Y$  sont **différents** si il existe un élément présent dans l'un mais pas dans l'autre -  $X$  et  $Y$  sont **disjoints** s'ils n'ont aucun élément commun :  $X \cap Y = \emptyset$  (disjoint est "plus fort" que différent)

*Exemples.* -  $\{0, 1\} \cup \{1, 2, 3\} = \{0, 1, 2, 3\}$  -  $\{-0, 2, 4, \dots\} \cup \{1, 3, 5, \dots\} = \mathbb{N}$  -  $\{0, 1\} \cap \{1, 2, 3\} = \{1\}$  : ils sont différents mais pas disjoints -  $\{0, 2, 4, \dots\} \cap \{1, 3, 5, \dots\} = \emptyset$  : ils sont disjoints

**Propriété.**  $\text{card}(X \cup Y) + \text{card}(X \cap Y) = \text{card } X + \text{card } Y$

Diagramme de Venn avec union et intersection

D'où  $\text{card } X \cup Y \leq \text{card } X + \text{card } Y$  et : on a l'égalité ssi  $X$  et  $Y$  sont disjoints. Dans ce cas on note l'union disjointe  $X \cup + Y$  ou  $X + Y$  alors  $\boxed{\text{card}(X + Y) = \text{card}(X) + \text{card}(Y)}$

### Différence

$$X \setminus Y = \{x \in X | x \notin Y\}$$

$$\text{card}(X \setminus Y) \leq \text{card } X$$

### Partition

**Définition.** Soit un ensemble  $X$  (fini ou non). On appelle **partition finie** de  $X$  :  $n$  sous-ensembles  $X_1, X_2, \dots, X_n$  deux à deux disjoints (= **exclusivité**) et dont l'union forme  $X$ .

Autrement dit : tout élément de  $X$  fait partie d'un  $X_i$  et d'un seul

*Exemple.*  $\mathbb{N}$  est partitionné en nombres pairs et nombres impairs.

$$X = X_1 + X_2 + \dots + X_n = \sum_{i=1}^n X_i$$

$$X = \bigcup_{i=1}^n X_i \text{ et } \forall i, j \text{ disjoints } X_i \cap X_j = \emptyset$$

*Exemples.* - L'ensemble des élèves d'INFO3 rangés par année de naissance. -  $\mathbb{Z}$  est partitionné en 5 parties  $(\mathbb{Z}/5\mathbb{Z})$  :  $\{5k | k \in \mathbb{Z}\}, \{5k + 1 | k \in \mathbb{Z}\}, \{5k + 2 | k \in \mathbb{Z}\}, \{5k + 3 | k \in \mathbb{Z}\}, \{5k + 4 | k \in \mathbb{Z}\}$

## Complémentaire

Soient  $X$  et  $Y$  deux ensembles tels que  $Y \subseteq X$ . Alors  $X - Y$  ou  $C_x Y$  est le supplémentaire de  $Y$  dans  $X$ .  $\{x \in X | x \notin Y\}$  : cas particulier de différence

**Alors :** -  $Y$  et  $X - Y$  forment 2 partitions de  $X$  :

$$X = (X - Y) + Y$$

$$\text{- } \text{card}X = \text{card}(X - Y) + \text{card}Y - \boxed{\text{card}(X - Y) = \text{card}X + \text{card}Y} \text{ si } Y \subseteq X$$

## c) Ensemble d'ensembles, n-uplets

Soient  $X_1, X_2, \dots, X_n$  des ensembles.

Le **produit cartésien** de ces ensembles, noté  $X_1 \times X_2 \times \dots \times X_n$  ou  $\prod_{i=1}^n X_i$  est l'ensemble des n-uplets de la forme  $(x_1, x_2, \dots, x_n)$  avec  $x_1 \in X_1, x_2 \in X_2, \dots, x_n \in X_n$

En informatique ce sont les **struct**, **enregistrement**, **tuple**, etc.

*Exemple.* -  $\{1, 2\} \times \{a, b, c\} = \{(1, a), (2, a), (1, b), (1, c), (2, b), (2, c)\}$  -  $\{1, 2\} \times \{1, 2, 3\} = \{(1, 1), (1, 2), (2, 1), (1, 3), (2, 2), (2, 3)\}$  ;  $(1, 2) \neq (2, 1)$  !

**Attention** l'ordre d'un n-uplet est important !

**Propriété.** Si  $X_1, \dots, X_n$  sont finis, alors :

$$\boxed{\text{card}(X_1 \times X_2 \dots X_n) = \text{card}(X_1) \times \text{card}(X_2) \times \dots \times \text{card}(X_n)}$$

**Cas particulier.** Si  $X_1 = X_2 = \dots = X_n = X$  alors on note  $X^n = X \times X \times \dots \times X$  et  $\text{card}(X^n) = \text{card}(X)^n$

*Exemple.*  $\mathbb{R}^n$  ; les mots de longueur  $n$  sur l'alphabet  $X$  "correspondent" exactement au éléments de  $X^n$

## Ensemble des parties

**Définition.** Si  $X$  est un ensemble, on note  $\mathcal{P}(X)$  l'ensemble des parties (ou sous-ensembles) de  $X$ , autrement dit tous les ensembles contenus dans  $X$ .

$$A \subseteq X \text{ ssi } A \in \mathcal{P}(X)$$

*Exemple.*  $X = \{a, b, c\}, \mathcal{P}(x) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{b, c\}, \{a, c\}, \{a, b, c\}\}$

*Remarque.* Contrairement aux mots, dans une partie il n'y a pas de répétition ni de notion d'ordre (il n'a pas d'importance)

**Propriété.**

$$\boxed{\text{card}(\mathcal{P}(X)) = 2^{\text{card}X}}$$

*Démonstration.* Si  $X = \{x_1, x_2, \dots, x_n\}$  une partie  $Y$  de  $X$  correspond à un mot binaire  $b_1 b_2 \dots b_i \dots b_n$  avec  $b_i = \begin{cases} 0 & \text{si } x_i \notin Y \\ 1 & \text{si } x_i \in Y \end{cases}$ , il y a autant de parties de  $X$  que de mots binaires de longueur  $n$  :  $(\text{card}\{0, 1\})^n = 2^n = 2^{\text{card}X}$

## Ensembles infinis

Parmi les ensembles **infinis**, on peut distinguer les ensembles **dénombrables**, c'est-à-dire ceux pour lesquels on peut **énumérer** les éléments (les numéroter  $0, 1, 2, \dots$ ) > Autrement dit faire une correspondance entre les entiers naturels et l'ensemble en question

*Exemple.*  $\mathbb{N}$  et  $\mathbb{Z}$  sont dénombrables mais pas  $\mathbb{R}$  ni  $[0, 1]$

## 4) Monoïdes

On appelle **monoïde** un ensemble  $X$  lorsque pour tous éléments  $x, y, z \in X$  : -  $X$  est pourvu d'une opération binaire interne ou **loi interne**  $\square : x \square y \in X$  -  $\square$  soit **associative** : si  $x, y, z \in X$  alors  $(x \square y) \square z = x \square (y \square z) = x \square y \square z$  -  $X$  possède un **élément neutre**  $e : e \square x = x = x \square e$

*Exemples.* -  $(\mathbb{R}, +, 0)$  -  $(\mathbb{R}, \times, 1)$  -  $(\mathbb{N}, +, 0)$  (**n'est pas un groupe !**) -  $(\mathbb{N}, \times, 1)$  -  $(A^*, \cdot, \epsilon)$  -  $(M_{n,n}(\mathbb{R}), \times, I_n)$  -  $(\mathcal{P}(X), \cap, X)$  -  $(\mathcal{P}(X), \cup, \emptyset)$

*Remarque.* Un groupe est toujours un monoïde, mais l'inverse n'est pas vrai.

**Propriété.** Dans un monoïde  $e$  est **unique**.

*Démonstration.* Supposons en effet que  $e$  et  $e'$  soient tous deux neutres, alors :

$$e = e \square e' = e'$$

**Homomorphisme de monoïdes.** Soient  $(X, \square, e)$  et  $(X', \square', e')$ , on dit que  $f : X \rightarrow X'$  est un **homomorphisme de monoïdes** si  $\begin{cases} f(e) = e' \\ f(x \square y) = f(x) \square' f(y) \end{cases}$  (il faut que les deux conditions soient respectées !)

*Exemples.* -  $\exp : (\mathbb{R}, +, 0) \rightarrow (\mathbb{R}_+^*, \times, 1)$  avec  $e^0 = 1$  et  $e^{x+y} = e^x \times e^y$  -  $\ln$  est l'homomorphisme de monoïdes réciproque -  $\lg : (A^x, \cdot, \epsilon) \rightarrow (\mathbb{N}, +, 0)$  avec  $\lg \epsilon = 0$  et  $\lg(u \cdot v) = \lg u + \lg v$  - complémentaire : -  $(\mathcal{P}(X), \cap, X) \rightarrow (\mathcal{P}(X), \cup, \emptyset)$  avec  $\overline{\overline{X}} = \emptyset$  et  $\overline{A \cap B} = \overline{A} \cup \overline{B}$  -  $(\mathcal{P}(X), \cup, \emptyset) \rightarrow (\mathcal{P}(X), \cap, X)$  -  $\text{card} : (\mathcal{P}(X), \cup, \emptyset) \rightarrow (\mathbb{N}, +, 0)$  n'est **pas** un homomorphisme de monoïdes car  $\text{card } \emptyset = 0$  mais  $\text{card}(A \cup B) \neq \text{card } A + \text{card } B$  (dès que  $A$  et  $B$  ne sont pas disjoints)

## 5) Systèmes de numération

**Définition.** Une **base de numération** est un entier  $b \geq 2$  et un symbole pour chaque valeur de  $0$  à  $b - 1$

*Exemples.* - La base 10 usuelle - La base 2 avec  $\{0, 1\}$  - La base 16 avec  $\{0, 1, \dots, 9, A, B, C, D, E, F\}$  - La base 60 (date des Mésopotamiens en -4000) est encore utilisée pour les minutes et secondes - La base 256 où on représente un chiffre par un couple d'hexadécimaux ( $16 \times 16 = 256$ ) est utilisée pour représenter des couleurs en informatique - La base 64 avec  $\{A \dots Z a \dots z 0 \dots 9 + /\}$  avec  $A'' = "0$ ,  $a'' = "26$  et  $0'' = "52$

**Définition.** L'**écriture en base  $b$  d'un entier  $n$**  est un mot sur l'alphabet des chiffres  $x_k \dots x_0$  tel

$$\text{que } \begin{cases} x_k \neq 0 \\ \sum_{i=0}^k x_i b^i = n \end{cases}$$

**Attention.** Différencier les symboles (écrits) de la valeur (entière) : l'entier qui vaut 7 en base 10 s'écrit 7, en base 2 il s'écrit 111 et en base 1 il s'écrit IIIIII.

**Propriété.** L'écriture en base  $b$  d'un entier  $n$  existe toujours et elle est unique

*Notation.* On écrit  $(x_k \dots x_0)_b$  pour noter la base



**Définition.** Si  $b^k \leq n < b^{k+1}$  alors **la taille de  $n$  en base  $b$**  est le nombre de chiffres qu'il faut pour l'écrire en base  $b$ , ici  $k + 1$ .

*Démonstration.* Chaque  $x_i$  est compris entre 0 et  $b - 1$  et  $x_k \geq 1$ .

D'où  $0 + 1.b^k \leq \underbrace{\sum_{i=0}^k x_i b^i}_{\leq \sum_{i=0}^k (b-1)b^i} \leq \sum_{i=0}^k (b-1)b^i$  ( $= b^{k+1} - 1 < b^{k+1}$ ) or,

$$\begin{aligned} \sum_{i=0}^k (b-1)b^i &= (b-1) + (b^2 - b) + (b^3 - b^2) + \dots + (b^{k+1} - b^k) \\ &= b^{k+1} - 1 \text{ (somme télescopique)} \end{aligned}$$

*Rappel.*  $\sum_{i=0}^k b^i = \frac{b^{k+1}-1}{b-1}$

On appelle  $\log_b$  (logarithme en base  $b$ ) une fonction croissante sur  $\mathbb{R}_+^*$  telle que  $\log_b(b^k) = k$ .

On peut la définir par  $\log_b x = \frac{\ln x}{\ln b}$ .

Alors immédiatement, si  $n$  s'écrit sur  $k + 1$  chiffres en base  $b$ , alors  $k = \lfloor \log_b n \rfloor$ .

Donc la taille de  $n$  en base  $b$  est  $\boxed{1 + \lfloor \log_b n \rfloor}$ .

*Exemple.* On veut écrire le nombre d'humains sur Terre (environ 7 milliards  $= 7 \times 10^9$ ) sur des bits, cherchons de combien de bits on a besoin :

$$\log_2(7 \times 10^9) = \log_2(7) + \log_2(10^9) \approx 3 + 30 (= 33)$$