Théorie des codes

Olivier Bernard

17 Décembre 2009

Résumé

La théorie des codes a vu le jour dans le giron de la théorie de l'information, introduite par Shannon vers 1950. Intuitivement, l'objet de la théorie des codes est l'étude des propriétés des factorisations des mots en une séquence de mots d'un ensemble donné. La théorie des codes (et notamment des codes correcteurs) est cruciale à notre époque où les flux de données chiffrées sont omniprésents.

Nous nous intéresserons ici à quelques propriétés basiques des codes.
Après une introduction formelle aux notions de code sur un alphabet et d'ensemble complet, nous ferons une première approche des liens entre codes complets et codes maximaux. Enfin, l'introduction des mesures permettra de renforcer ces liens et de donner des critères efficaces de caractérisation.

1 Premières définitions

1.1 Codes

Soit A un alphabet et X un sous-ensemble de A^*. Intuitivement, X est un code si tout mot de X^+ a au plus une factorisation en produit de mots de X. En d'autres termes :

Définition 1.1 *$X \subsetneq A^*$ est un* code *si :*

$$\begin{cases} \forall n, m \in \mathbb{N} \setminus \{0\} \\ \forall (x_i)_{i\in[\![1,n]\!]}, (x'_j)_{j\in[\![1,m]\!]} \in X, \end{cases} \quad x_1 \dots x_n = x'_1 \dots x'_m \quad \Longrightarrow \begin{cases} n = m \\ \forall i \in [\![1, n]\!], \quad x_i = x'_i. \end{cases}$$

Remarques :

- Tout sous-ensemble d'un code est encore un code. En particulier, l'ensemble vide est un code.
- Pour tout code X, $\varepsilon \notin X$ [1]. En effet, la relation $\varepsilon = \varepsilon \cdot \varepsilon$ est incompatible avec la définition ci-dessus.

Exemple : Soit $A = \{a, b\}$. L'ensemble $X = \{a, ba\}$ est un code. Par contre, l'ensemble $X = \{ab, aba, aab\}$ n'est pas un code, car $aba \cdot ab = ab \cdot aba$.

1. ε désigne, et ce dans toute la suite également, le mot vide sur A.

Nous allons maintenant voir une classe importante de codes : les codes préfixes. Les codes préfixes sont les plus simples à construire, à partir d'automates déterministes par exemple, et permettent de résoudre une large classe de problèmes. Le lecteur intéressé se reportera fructueusement à [1, Chap. 2] pour un tour d'horizon sur le sujet. Rappelons tout d'abord la définition d'un ensemble préfixe :

Définition 1.2 *Soit $X \subsetneq A^*$. X est un ensemble* préfixe *si aucun élément de X n'est un facteur gauche propre d'un autre élément de X, ie :*

$$X \text{ préfixe} \quad \Longleftrightarrow \quad \forall x, x' \in X,\ (x \leq x' \Rightarrow x = x').$$

Remarque : Si X est préfixe et contient le mot vide ε, alors $X = \{\varepsilon\}$.

Proposition 1.3 *Tout ensemble préfixe $X \neq \{\varepsilon\}$ est un code.*

Preuve : Si X n'est pas un code, il existe deux décompositions $u = x_1 \dots x_n = x'_1 \dots x'_m$ de u dans X^* comme dans la définition 1.1, telles que x_1, x'_1 sont non vides et $x_1 \neq x'_1$. Dans ce cas, $x_1 < x'_1$ ou $x'_1 < x_1$, ce qui est une contradiction à l'hypothèse que X est préfixe. □

Exemple : L'ensemble $X = a^*b$ est préfixe, c'est donc un code. Il est *infini* sur un alphabet fini.

1.2 Ensembles complets

Soit S un sous-ensemble (quelconque) de A^*.

Définition 1.4 *Un mot $m \in A^*$ est dit* complétable dans S *s'il existe $u, v \in A^*$ tels que $umv \in S$, ce qui est équivalent à demander $A^*mA^* \cap S \neq \emptyset$.*

L'ensemble des mots complétables dans S est noté $F(S) = (A^*)^{-1}S(A^*)^{-1}$. Il n'est pas difficile de vérifier que $A^* \setminus F(S)$ est un idéal bilatère de A^* [2].

Cette définition préliminaire permet de définir la notion de densité, qui conduit ensuite naturellement à la notion d'ensemble complet.

Définition 1.5 *On dit que S est* dense dans A^* *si tous les mots de A^* sont complétables dans S, ie $F(S) = A^*$. Dans le cas contraire, on dit que S est* maigre.

L'appellation « dense » se justifie de la façon suivante :
- S est dense $\Longleftrightarrow \forall m \in A^*, \quad A^*mA^* \cap S \neq \emptyset \Longleftrightarrow S$ rencontre tous les idéaux bilatères de A^*.

2. Un idéal bilatère d'un monoïde M est un sous-ensemble non-vide I de M tel que $MIM = I$.

– $\mathcal{O} = \{\emptyset\} \cup \{\text{idéaux bilatères de } A^*\}$ définit une topologie sur A^*. En effet, cet ensemble est stable par intersection finie et par union quelconque.

Dans ce cas, la définition 1.5 de la densité est équivalente à la définition topologique induite par la topologie $\mathcal{O}$.

L'intuition sous-jacente au terme « maigre » se trouve quant à elle dans la proposition suivante :

Proposition 1.6 *Soit P, Q et R des sous-ensembles de A^*.*

1. *$P \cup Q$ est maigre $\iff$ P et Q sont maigres.*
2. *Si R est dense et P est maigre, alors l'ensemble $R \setminus P$ est encore dense.*

Preuve : La première propriété implique la seconde. En effet, si $R \setminus P$ est maigre, $(R \setminus P) \cup P = R$ est maigre, ce qui est une contradiction.
Supposons maintenant que P et Q sont maigres. Alors il existe un mot m et un mot n sur A tels que $A^*mA^* \cap P = \emptyset = A^*nA^* \cap Q$. Ainsi,

$$A^*mnA^* \cap (P \cup Q) \subseteq (A^*mA^* \cap P) \cup (A^*nA^* \cap Q) = \emptyset,$$

ce qui termine ce côté de la preuve.
Réciproquement, s'il existe $x \in A^*$ tel que $A^*xA^* \cap (P \cup Q) = \emptyset$, alors $A^*xA^* \cap P \subseteq \emptyset$, et symétriquement en Q. □

Voici maintenant, comme annoncé, la définition d'un ensemble complet :

Définition 1.7 *On dit que l'ensemble $S \subseteq A^*$ est* complet *si le sous-monoïde de A^* engendré par S est dense, ie si $F(S^*) = A^*$.*

Remarques :

– Si S est un sous-monoïde de A^*, on a en fait l'équivalence S dense $\iff$ S complet. Attention cependant, si l'implication S dense $\implies$ S complet est vraie, l'affirmation inverse est fausse en général : l'ensemble préfixe $X = a^*b$ sur $A = \{a, b\}$ n'est pas dense [3], mais il est bel et bien complet.
– Tout ensemble qui contient un sous-ensemble complet est, *a fortiori*, lui-même complet.

Dans tout ce qui précède, il est possible de remplacer A^* par un monoïde M quelconque, sans complications supplémentaires. Avant d'aborder la section suivante, qui mettra ces différentes notions en rapport, nous sommes contraints de donner un résultat aussi technique que malheureusement indispensable pour la suite.

3. Par exemple, *aba* n'est pas complétable dans X.

Lemme 1.8 *Soit $X \subsetneq A^*$ un ensemble maigre complet.*
Soit w un mot de A^* incomplétable *dans X*[4]*. On note $\mathcal{G}$ (resp. $\mathcal{D}$) l'ensemble (fini) des facteurs gauches (resp. droits) de w. Alors :*

$$A^* = \bigcup_{\substack{d\in\mathcal{D}\\ g\in\mathcal{G}}} d^{-1}X^*g^{-1} = \mathcal{D}^{-1}X^*\mathcal{G}^{-1}.$$

En dépit de son aspect austère, ce lemme a le mérite d'exprimer l'ensemble des mots sur A comme une réunion *finie* d'ensembles construits à partir du monoïde engendré par un code maigre (mais complet) et d'un seul mot incomplétable.

Preuve : Soit $z \in A^*$. Par densité de X^*, il existe deux mots u et v de A^* tels que $u \cdot wzw \cdot v \in X^*$. Le mot w est incomplétable dans X, ce qui implique que toute décomposition de $uwzwv$ en mots de X scinde les deux occurrences de w en deux mots non vides. On peut aussi supposer u et v de longueur minimale, de sorte que $uwzwv$ se décompose en mots de X sous la forme

$$uwzwv = (ug')(dzg)(d'v), \qquad \text{où } \begin{cases} w = g'd = gd' \\ ug', d'v \in X \text{ et } dzg \in X^* \end{cases}$$

Ainsi, $z \in d^{-1}X^*g^{-1}$. On obtient ainsi une inclusion, et l'inclusion réciproque est évidente. □

1.3 Mots sans bords

Cette section se termine sur une partie technique sans rapport direct avec les codes, mais dont le résultat est essentiel pour la preuve du théorème 2.6.

Définition 1.9 *Un mot $w \in A^*$ est dit* sans bords *si aucun des facteurs gauches propres non vides de w n'en est aussi un facteur droit (propre). En d'autre termes, dans ce cas :*

$$w \in uA^* \cap A^*u \quad \Longrightarrow \quad u = \varepsilon \text{ ou } u = w.$$

Lemme 1.10 *Soit A un alphabet tel que $\#A \geq 2$. Pour tout mot $u \in A^+$, il existe un mot $v \in A^*$ tel que uv est sans bords.*

Preuve : Soit a la première lettre de u, et $b \in A \setminus \{a\}$. Le mot $v = ab^{|u|}$ convient.
En effet, un facteur gauche non vide t de uv commence par la lettre a. Si t est aussi un facteur droit de uv, $|t| > |u|$ et t termine par $|u|$ instances de b. Alors t s'écrit $sab^{|u|}$, $s \in A^*$. Or t est un facteur gauche de uv, et un argument de longueur permet alors d'écrire $t = uab^{|s|}$. Ceci implique $|s| = |u|$, et donc $t = uv$. □

4. Un tel mot existe, car X est maigre.

2 Codes complets et maximaux

Soit A un alphabet.

Le but de cette section est de préciser, dans le cas des codes sur A, les liens qui existent entre la notion de maximalité, introduite au paragraphe suivant, et celle de complétude. Il s'agit en fait d'essayer de remplacer une propriété d'optimum par une propriété combinatoire, plus simple à appréhender.

2.1 Codes maximaux

Nous avons en fait déjà pu remarquer que tout sous-ensemble d'un code est encore un code. Ceci invite à s'intéresser de plus prêt à la structure des codes maximaux. Intuitivement, un code est maximal s'il n'est inclus strictement dans aucun autre. C'est formellement la définition naturelle suivante :

Définition 2.1 *Un code X est dit* maximal sur A *si pour tout code X' sur A, on a l'implication suivante :*

$$X \subseteq X' \subsetneq A^* \quad \Longrightarrow \quad X = X'.$$

Remarque : Le choix de l'alphabet a ici son importance. En effet, prenons B un alphabet tel que $A \subsetneq B$ et X un code maximal sur A. Alors X est *a fortiori* un code sur B, mais il n'est *pas* maximal ($\forall b \in B \setminus A, \quad X \cup \{b\}$ est encore un code).

L'important théorème suivante justifie l'intérêt porté aux codes maximaux.

Théorème 2.2 *Tout code X sur A est inclus dans un code maximal sur A.*

Preuve : Le schéma de la preuve est un argument classique à base de lemme de Zorn.
Soit $\mathcal{F} = \{W \text{ tq } W \text{ code sur } A,\ X \subseteq W\}$, ordonné par l'inclusion.

- $\mathcal{F}$ est trivialement non vide ($X \in \mathcal{F}$).
- $\mathcal{F}$ est inductif : soit $\mathcal{C}$ une chaîne d'inclusions de $\mathcal{F}$. $\widehat{W} = \bigcup_{W \in \mathcal{C}} W$ est une borne supérieure de $\mathcal{C}$. Il reste à montrer que $\widehat{W}$ est un code.
  Soit $n, m \in \mathbb{N} \setminus \{0\}, (w_i)_{i \in [\![1,n]\!]}, (w'_j)_{j \in [\![1,m]\!]} \in \widehat{W}$ tels que $w_1 \dots w_n = w'_1 \dots w'_m$. Par définition de $\mathcal{C}$ et de $\widehat{W}$, il existe $Y \in \mathcal{C}$ qui contient tous les w_i, w'_j. Alors, Y est un code, et la définition 1.1 assure $n = m$ et $\forall i \in [\![1, n]\!], \quad w_i = w'_i$.

Donc $\mathcal{F}$ est un ensemble ordonné inductif non vide, qui admet, d'après le lemme de Zorn, un élément maximal. □

Remarque : Ce résultat est faux si l'on se restreint aux codes finis. Il est en effet possible de construire des codes finis qui ne sont inclus dans aucun code maximal *fini* (cf. [1, Chap. I, Ex : 5.6]).

2.2 Codes complets

Le premier théorème de ce paragraphe est à mettre en rapport avec le théorème 2.2.

Théorème 2.3 *Tout code X sur un alphabet A est inclus dans un code* complet *sur A.*

La preuve fournit une construction explicite, bien qu'un rien fastidieuse.

Preuve : Soit $X \subsetneq A^+$ un code, et soit $y \in A^+$ un mot sans bords, tel que y n'est pas complétable dans X.
On pose $U = A^* \setminus X^* \setminus A^*yA^* \neq \emptyset$. On va montrer que $Y = X \cup y(Uy)^*$ est un code complet.

Lemme 2.4 *Si $V = A^* \setminus A^*yA^*$, l'ensemble $Vy \neq \{\varepsilon\}$ est un code* préfixe.

Preuve : Supposons, pour v, $v' \in V$, que $vy < v'y$. Comme y est sans bords, il est facile de voir que $vy \leq v'$. Ainsi, v' appartient à $A^*yA^* \cap V = \emptyset$, ce qui est impossible. On a donc bien $vy \leq v'y$ implique $vy = v'y$, ce qui termine la preuve. □

Lemme 2.5 *Y est un code.*

Preuve : C'est en fait le point le plus ardu.
Supposons que Y n'est pas un code. Il existe alors alors deux décompositions $y_1 \dots y_n = y'_1 \dots y'_m$ dans Y^* comme dans la définition 1.1 telles que $y_1 \neq y'_1$.
Comme X est un code, il existe au moins un de ces facteurs qui n'est pas dans X. Disons qu'il s'agit de y_p, où p est le plus petit indice tel que $y_p \in Y \setminus X = y(Uy)^*$. Le mot y_p n'est pas complétable en un mot de X (sinon il existe u, $v \in A^*$ tels que $uy_pv \in X \cap uy(Uy)^* \subsetneq X \cap A^*yA^*$ et y serait complétable dans X). Il existe donc un indice q minimal tel que y'_q est lui aussi dans $Y \setminus X$. Ainsi, $y_1 \dots y_{p-1}y$ et $y'_1 \dots y'_{q-1}y$ sont dans $X^*y \subsetneq Vy$, et vu que l'on a montré que Vy est un ensemble préfixe, on a $y_1 \dots y_{p-1} = y_1 \dots y_{q-1}$.
Les hypothèses X est un code et $y_1 \neq y'_1$ impliquent alors y_1, $y'_1 \in y(Uy)^*$. On écrit maintenant

$$y_1 = yu_1y \dots yu_ky, \text{ et } y'_1 = yu'_1y \dots yu'_ly, \text{ où les } (u_i)_{i\in[\![1,k]\!]},\ (u'_j)_{j\in[\![1,l]\!]} \in U \subsetneq V.$$

Supposons que $y_1 < y'_1$. Comme Vy est préfixe, pour tout $i \in [\![1,k]\!]$, $u_i = u'_i$. Le facteur restant dans y'_1 est alors $t = u'_{k+1}y \dots yu'_l$.
On termine alors la preuve en écrivant $y_2 \dots y_n = ty'_2 \dots y'_m$. Le mot y est un facteur de t, il existe donc un r minimal tel que $y_r \in y(Uy)^*$. Mais à nouveau, l'utilisation du fait que Vy est préfixe, combinée avec la remarque que $y_2 \dots y_{r-1}y$ et $u'_{k+1}y$ sont des facteurs gauches du même mot, donne la relation $y_2 \dots y_{r-1} = u'_{k+1}$. Donc, par hypothèse de la minimalité de r, $u'_{k+1} \in X^* \cap U = \emptyset$, ce qui est absurde.

□

Il reste à montrer que Y est complet. Soit $w \in A^*$, w s'écrit $w = v_1yv_2y\cdot\cdots\cdot yv_n$, où $n \in \mathbb{N}\setminus\{0\}$, et où les $v_i \in A^*\setminus A^*yA^* = V$ sont éventuellement vides. Mais dans ce cas, $ywy \in Y^*$, ce qui termine la preuve du théorème. □

Les deux paragraphes précédents ont permis de voir que tout code est inclus dans un code maximal, et que tout code peut être complété. Le théorème fondamental suivant met en évidence le lien soupçonné.

Théorème 2.6 *Tout code maximal est complet.*

Preuve : Soit $X \subsetneq A^*$ un code non complet. On montre que X n'est pas maximal.

- Si $\#A = 1$, les codes non vides sur A sont les singletons, qui sont évidemment complets. Ainsi, $X = \emptyset$, qui n'est pas maximal.
- Si $\#A \geq 2$, comme X n'est pas complet, il existe un mot u non vide sur A non complétable sur X^*. D'après le lemme 1.10, il existe $v \in A^*$ tel que $y = uv$ est sans bords. On vérifie que y n'est pas complétable sur X^* : si tel est le cas, u l'est aussi, contradiction.
  On peut alors appliquer le théorème 2.3 pour affirmer que $X \cup y(Uy)^*$ est un code qui contient strictement X.

Donc X n'est pas maximal, ce qui termine la preuve. □

Remarque : La réciproque est fausse en général. Pour un code X dense maximal, et pour tout élément $x \in X$, le code $X\setminus\{x\}$ est toujours dense, d'après la proposition 1.6, mais bien sûr, il n'est plus très maximal.

Néanmoins, l'affirmation inverse est valable pour les codes *maigres*. Il va être nécessaire, pour établir ce résultat, de formaliser l'intuition qu'un code n'a que peu de mots. C'est l'objet de la section suivante.

3 Mesures

On cherche à construire une mesure sur A^*. Il est naturel d'imaginer, par exemple, que les lettres d'un mot suivent des lois de probabilités indépendantes identiquement distribuées. Ceci amène au premier paragraphe.

3.1 D'une distribution de probabilité...

La définition suivante formalise l'idée développée ci-dessus.

Définition 3.1 *Une* distribution de Bernoulli *sur A^* est un morphisme de monoïdes $\pi : A^* \longrightarrow (\mathbb{R}_+, \times)$ tel que $\sum_{a\in A} \pi(a) = 1$. Elle est dite* positive *si $\forall a \in A,\ \pi(a) \neq 0$.*

Remarques :

- Le lecteur prendra bien garde de considérer $\mathbb{R}_+$ comme un monoïde *multiplicatif*.
- Il découle de la définition que $\forall a \in A$, $\pi(a) = \pi(a \cdot \varepsilon) = \pi(a)\pi(\varepsilon)$. Ceci étant vrai pour tout a, la condition de sommation impose $\pi(\varepsilon) = 1$.

Il est alors aussi judicieux que tentant de poser, $\forall L \subseteq A^*,\ \pi(L) = \sum_{l \in L} \pi(l)$. Par abus, la fonction ainsi obtenue sur $\mathfrak{P}(A^*)$ est encore notée π.

Proposition 3.2 *Soit π une distribution de Bernoulli sur A^*.*
L'application $\pi : \mathfrak{P}(A^*) \longrightarrow \mathbb{R}_+ \cup \{\infty\}$ est une mesure positive *sur A^*, ie :*

1. $\forall L \subseteq A^*, \pi(L) \geq 0 \quad et \quad \pi(\emptyset) = 0$;
2. $\forall (E_n)_{n\in\mathbb{N}}$ *tq les* $E_n \subseteq A^*$ *sont deux à deux disjoints,* $\pi(\cup E_n) = \sum \pi(E_n)$.

Preuve : La première propriété est immédiate, par définition de π, et la seconde propriété implique, comme $\pi(A) = 1 < \infty$, que $\pi(\emptyset) = 0$.
Soit $(E_n)_{n\in\mathbb{N}}$ une suite de sous-ensembles de A^* deux à deux disjoints.

$$\pi(\cup E_n) = \sum_{l \in \cup E_n} \pi(l) = \sum_{n\geq 0} \left(\sum_{l \in E_n} \pi(l) \right) = \sum_{n \geq 0} \pi(E_n).$$

□

Remarques :

- Ainsi $\pi(A^*) = \sum_n \pi(A^n) = \sum_n 1 = \infty$.
- En général, pour une famille quelconque $(E_n)_{n\in\mathbb{N}}$, on a $\pi(\cup E_n) \leq \sum \pi(E_n)$.

Voici maintenant deux propriétés utiles pour calculer la mesure d'un ensemble. Pour $L \subseteq A^*$ et pour tout entier $n \geq 0$, on note $L_n = L \cap A_{\leq n} = \{u \in L \text{ tq } |u| \leq n\}$.

Proposition 3.3 *Soit π une distribution de Bernoulli sur A^*, et $L, M \subseteq A^*$. Alors :*

1. $\pi(L) = \lim_{n\to\infty} \pi(L_n)$;
2. $\pi(LM) \leq \pi(L)\pi(M)$.

Preuve : Pour tout entier $n \geq 0$, $L \cap A_{\leq n} \subset L \cap A_{\leq n+1}$. Par les propriétés des mesures :

$$\pi(L) = \pi\left(\bigcup_{n\geq 0} L_n \right) = \lim_{n\to\infty} \pi(L_n).$$

Par ailleurs, par la remarque précédente,

$$\pi(LM) = \pi\left(\bigcup_{l\in L} \bigcup_{m\in M} \{lm\} \right) \leq \sum_{l\in L} \sum_{m\in M} \pi(l)\pi(m) = \pi(L)\pi(M).$$

□

3.2 ... à différentes caractérisations des codes

Soit $X \subsetneq A^+$ et π une distribution de Bernoulli sur A^*.
Le paragraphe précédent a en particulier attiré l'attention sur le fait que

$$\pi(X^*) = \pi\left(\bigcup_{n\in\mathbb{N}} X^n\right) \leq \sum_{n\in\mathbb{N}} \pi(X^n) \leq \sum_{n\in\mathbb{N}} \pi^n(X).$$

En fait, si X est un code, il y a partout égalité, et moyennant quelques précautions, cela est même une condition suffisante. Notons déjà que par définition d'un code, les X^n sont deux à deux disjoints, et ainsi, par les propriétés des mesures :

$$(X \text{ code}) \qquad \pi(X^*) = \sum_{n\in\mathbb{N}} \pi(X^n).$$

Proposition 3.4

1. *Si X est un code, alors $\forall n \in \mathbb{N}$, $\pi(X^n) = \pi^n(X)$.*
2. *Si π est* positive *et si $\pi(X) < \infty$, alors :*

$$(\forall n \in \mathbb{N}, \quad \pi(X^n) = \pi^n(X)) \quad \Longrightarrow \quad X \textit{ est un code.}$$

Remarque : En particulier, le premier point implique que $\pi(X^*) < \infty \iff \pi(X) < 1$, dans le cas où X est un code.

Preuve : Soit $\chi_n = X \times X \times \ldots X$ le produit cartésien de n instances de X.

1. Alors l'application :

$$\begin{array}{rccc} \varphi & : & \chi_n & \longrightarrow & X^n \\ & & (x_1, \ldots, x_n) & \longmapsto & x_1 \cdot \cdots \cdot x_n \end{array}$$

   est clairement surjective, et injective par définition d'un code. Ainsi :

$$\begin{array}{rcll} \pi^n(X) & = & \left(\displaystyle\sum_{x_i\in X} \pi(x_i)\right)^n & = & \displaystyle\sum_{(x_1,\ldots,x_n)\in\chi_n} \pi(x_1)\ldots\pi(x_n) \\ & & & = & \displaystyle\sum_{x\in X^n} \pi(x) \quad \text{par la bijection } \varphi \\ & & & = & \pi(X^n) \end{array}$$

2. Réciproquement, supposons que π est positive, que $\pi(X)$ est finie et que, pout tout $n \geq 0$, $\pi(X^n) = \pi^n(X)$. Si X n'est pas un code, il existe un mot $u \in X^*$ qui a deux décompositions différentes, par exemple $u = x_1 \ldots x_n = x'_1 \ldots x'_m$, pour des éléments de X et des entiers n, m non nuls. Dans X^{n+m}, uu a deux décompositions différentes au moins, à savoir $uu = x_1 \ldots x_n \cdot x'_1 \ldots x'_m = x'_1 \ldots x'_m \cdot x_1 \ldots x_n$. Mais,

si les deux décompositions sont présentes dans le produit cartésien χ_{n+m}, une seule instance est décomptée dans la mesure de X^{n+m}, ie formellement :

$$\pi^{n+m}(X) = \sum_{(y_1,\ldots,y_k)\in\chi_{n+m}} \pi(y_1)\ldots\pi(y_k) \quad \geq \quad \pi(X^{n+m}) + \pi(uu).$$

Ceci, par finitude de la mesure de X, et donc de celle de X^{n+m}, implique que $0 < \pi(uu) \leq 0$, par l'hypothèse que π est positive, ce qui est absurde.

□

La proposition suivante, quant à elle, formalise l'idée que les codes n'ont que peu de mots. Elle est surtout utile en pratique pour montrer qu'un ensemble donné n'est pas un code.

Proposition 3.5 *Soit X est un code sur A.*
Pour toute *distribution de Bernoulli π sur A^*, $\pi(X) \leq 1$.*

Preuve : Soit $X_k = X \cap A_{\leq k} = \{x \in X \text{ tq } |x| \leq k\}$, pour $k \geq 1$, et soit π une distribution de Bernoulli quelconque.
On a $X_k \subsetneq A \cup \cdots \cup A^k$, donc pour tout $n \geq 1$, $X_k^n \subsetneq A \cup \cdots \cup A^{nk}$. Ainsi,

$$\pi(X_k^n) \leq \sum_{i=1}^{nk} \pi(A^i) = nk.$$

Supposons maintenant que $\pi(X_k) > 1$, ie $\pi(X_k) = 1 + \varepsilon$, avec $\varepsilon > 0$. L'affirmation précédente implique que pour tout $n \geq 1$, $(1+\varepsilon)^n \leq nk$. Or $\lim_{n\to\infty} \dfrac{(1+\varepsilon)^n}{n} = \infty$, ce qui est une contradiction.
Donc, pour tout $k \geq 1$, $\pi(X_k) \leq 1$ et $\pi(X) = \lim_{k\to\infty} \pi(X_k) \leq 1$. □

Remarque : L'hypothèse « X est un code » est cruciale dans cette preuve : c'est grâce à elle que l'on peut écrire $\pi(X_k^n) = \pi^n(X_k)$, et trouver la contradiction recherchée.

Ainsi, pour montrer qu'un ensemble n'est pas un code, il suffit d'exhiber une distribution de Bernoulli sur A^* telle que $\pi(X) > 1$.

Attention cependant, cette condition n'est pas suffisante : il est possible de générer un sous-ensemble de A^* de mesure toujours inférieure à 1, et qui n'est néanmoins pas un code, comme le montre l'exemple ci-dessous.

Exemple : On a déjà vu que $X = \{ab, aba, aab\}$ n'est pas un code sur $A = \{a, b\}$. Si π est une distribution sur A^*, on peut poser $\pi(a) = p = 1 - \pi(b)$. Alors :

$$\pi(X) = p(1-p) + 2p^2(1-p) = p + p^2 - 2p^3.$$

Une brève étude de fonction montre que $\forall p \in [0,1], \pi(X) < 1$.

Ce paragraphe se termine par une caractérisation des codes maximaux.

Théorème 3.6 *Soit X un code sur A.*
S'il existe une distribution de Bernoulli positive *telle que $\pi(X) = 1$, alors X est maximal.*

Remarque : Il convient d'être très attentif aux hypothèses. D'une part, on suppose que X est un code (et non un ensemble quelconque), et d'autre part, on exclut des distributions qui ne seraient pas positives !
Par exemple, le code $X = \{a\}$ sur $A = \{a, b\}$ n'est pas maximal[5]. Pourtant si on pose $\pi(a) = 1, \pi(b) = 0$, π est une distribution *non positive*, et $\pi(X) = 1$.

Preuve : Soit π une distribution de Bernoulli sur A^* telle que $\pi(X) = 1$.
Si X n'est pas maximal, il existe un mot $y \in A^* \setminus X$ tel que $Y = X \cup \{y\}$ est un code. Alors, d'après le théorème 3.5 :

$$1 \geq \pi(Y) = \pi(X) + \pi(y) = 1 + \pi(y) \quad \Longrightarrow \pi(y) = 0,$$

ce qui est impossible. □

Jusqu'à présent, nous avons donc montré, pour X un code, la suite d'implications :

$$\exists \pi \text{ distribution de Bernoulli positive tq } \pi(X) = 1 \Longrightarrow X \text{ maximal } \Longrightarrow X \text{ complet.}$$

Dans le cas des codes maigres, il y a en fait équivalence, et c'est ce que se propose de montrer le dernier paragraphe de ce travail.

3.3 Codes maigres

La proposition ci-dessous rend en fait le résultat immédiat. Elle utilise le fait que la mesure de A^* est infinie, ainsi que le lemme 1.8 qui décompose A^* en une union finie d'ensembles :

Proposition 3.7 *Soit $X \subsetneq A^*$ un ensemble maigre et complet.*
Pour toute *distribution de Bernoulli positive sur A^*, $\pi(X) \geq 1$.*

Preuve : On reprend les notations du lemme 1.8. Comme $\pi(A^*) = \infty$, il existe un couple $(g, d) \in \mathcal{G} \times \mathcal{D}$ tel que $\pi(d^{-1}X^*g^{-1}) = \infty$.
Si on montre que $\pi(X^*) = \infty$, on aura, par la proposition 3.2 sur les propriétés des mesures :

$$\pi(X^*) = \infty \leq \sum_{n \in \mathbb{N}} \pi(X^n) \leq \sum_{n \in \mathbb{N}} \pi^n(X).$$

Et ceci implique $\pi(X) \geq 1$.

5. $X \subsetneq \{a, ba\}$, qui est un code, comme vu auparavant.

Mais, $d \cdot (d^{-1}X^*g^{-1}) \cdot g \subseteq X^*$, donc, par l'hypothèse que π est positive, on peut écrire $\pi(d)\pi(d^{-1}X^*g^{-1})\pi(g) = \infty \leq \pi(X^*)$, ce qu'on voulait. □

Remarque : Si on suppose de plus que X est un code (maigre et complet), alors la proposition 3.5 donne gratuitement, pour toute distribution de Bernoulli *positive*, $\pi(X) = 1$.

On peut combiner tous ces résultats dans les deux théorèmes suivants :

Théorème 3.8 *Soit X un code sur A. On a l'équivalence suivante :*

$$X \text{ complet} \iff X \text{ dense ou } X \text{ maximal}$$

Preuve : Si X est dense, il est complet, d'après la remarque qui suit la définition 1.7. Le théorème 2.6 affirme qu'un code maximal est complet.
Réciproquement, supposons que X est un code complet. Si X n'est pas dense (ie X est maigre), la remarque suivant la proposition 3.7 ci-dessus dit alors exactement que X est maximal. □

Théorème 3.9 *Soit X un code maigre sur A. Les affirmations suivantes sont équivalentes :*

1. *Pour toute distribution de Bernoulli positive π sur A^*, $\pi(X) = 1$.*

(1.') *Il existe une distribution de Bernoulli positive π sur A^* telle que $\pi(X) = 1$.*

2. *X est un code maximal.*
3. *X est un code complet.*

Preuve : (1) ⇒ (1') est laissée en exercice.
(1') ⇒ (2) C'est le théorème 3.6.
(2) ⇒ (3) C'est le théorème 2.6.
(3) ⇒ (1) C'est la proposition 3.7. □

Ce dernier théorème fournit donc des critères assez facilement appréhendables, pour un code maigre, qui permettent de montrer des notions *a priori* délicates comme la maximalité.

Table des matières

Références

[1] J. Berstel, D. Perrin : *Theory of codes*, Pure and applied mathematics (Academic Press) 1985, pp. 37-68.