RAIRO
INFORMATIQUE THÉORIQUE

MAX VINCENT

Construction de codes indécomposables

RAIRO – Informatique théorique, tome 19, n° 2 (1985), p. 165-178.

<http://www.numdam.org/item?id=ITA_1985__19_2_165_0>

R.A.I.R.O. Informatique théorique/Theoretical Informatics
(vol. 19, n° 2, 1985, p. 165 à 178)

CONSTRUCTION DE CODES INDÉCOMPOSABLES (*)

par Max VINCENT (1)

Communiqué par J. E. PIN

Résumé. — *La théorie des codes s'est développée à la suite des travaux de M. P. Schützenberger sur les langages formels, les codes apparaissant comme un objet fondamental de la combinatoire du monoïde libre. Nous construisons une nouvelle famille de sous-monoïdes libres, finiment engendrés, dont la base est asynchrone et indécomposable par des codes préfixes ou suffixes.*

Abstract. — *The theory of codes is issued from M. P. Schützenberger's works on formal languages; codes appear as a fundamental object of combinatorics on free monoïds. We build a new family of free submonoïds, finitely generated, the basis of which is asynchronous and undecomposable with prefix or suffix codes.*

1. INTRODUCTION

Les premieres études sur les codes datent des travaux de Shannon sur le codage des messages pour la communication (1950). Un deuxième type d'étude s'est développé à la suite des travaux de M. P.Schützenberger concernant la théorie des langages formels, les codes apparaissant ici comme bases des sous-monoïdes libres.

Les chercheurs s'intéressèrent d'abord aux codes préfixes, d'une part ils semblaient plus « naturels » (ils sont déchiffrables instantanément de gauche à droite) d'autre part tous les codes alors connus pouvaient s'obtenir par codages successifs au moyen de codes préfixes ou suffixes (correspondant au déchiffrage instantané de droite à gauche).

En 1972, Y. Cesari [4] produisait un code n'appartenant pas à cette classe ce qui posait le problème de la caractérisation des codes indécomposables.

Les méthodes d'études utilisant des représentations par des monoïdes finis d'applications ou de relations non ambigus ont permis d'obtenir des résultats importants. Néanmoins, pour la construction de familles de codes ayant des

(*) Reçu en décembre 1983, révisé en juillet 1984.

(1) Université Paul-Valéry, B.P. n° 5043, 34032 Montpellier Cedex.

propriétés particulières les chercheurs ont plutôt utilisé le langage des séries formelles qui permet de travailler directement sur l'expression du code et non sur ses représentations.

C'est grâce à cette méthode que J. M. Boë [1] produisait en 1978 des familles de codes indécomposables « synchronisants » et D. Perrin [10] en 1976 des familles de codes indécomposables « préfixes asynchrones ». Nous reprenons ici leurs traces pour construire des familles de codes indécomposables encore plus générales.

1.1. Monoïdes libres

Soit A un ensemble fini appelé alphabet; on appelle mot une suite finie d'éléments de A, la suite vide est notée 1. L'ensemble des mots sur A est muni par la concaténation d'une structure de monoïde. Ainsi défini, cet ensemble noté A^*, est le monoïde libre sur A.

On appelle longueur d'un mot de A^* le nombre de lettres noté $|m|$ qui le composent. Un facteur droit (resp. gauche) de m est un mot f de A^* tel que :

$$m = vf \qquad \text{avec} \quad v \text{ dans } A^* \text{ (resp. } m = fv).$$

Si $v \neq 1$ le facteur est dit propre.

Si P et Q sont deux parties de A^* nous poserons :

$$PQ = \{pq;\ p \text{ dans } P \text{ et } q \text{ dans Q}\}.$$

DÉFINITION : Si les mots d'un sous-monoïde P de A^* se factorisent de façon unique en produit de mots de son générateur minimal X, on dit que X est un code.

P est alors isomorphe au monoïde libre sur l'alphabet X; P est un sous-monoïde libre de A^* engendré par X, on le note X^*.

Définissons quelques classes importantes de codes : un code X est préfixe (resp. suffixe) s'il ne contient aucun de ses facteurs gauches (resp. droits) propres. Il est bipréfixe s'il est simultanément préfixe et suffixe.

1.2. Maximalité et complétude

DÉFINITION : Un code X sur A est maximal s'il n'est contenu dans aucun autre code sur A.

DÉFINITION : Un sous-monoïde P de A^* est complet s'il vérifie la condition :

$$\text{Pour tout } m \text{ de } A^* \text{ on a : } P \cap A^* m A^* \neq \varnothing.$$

PROPOSITION [12] : *Soit X un code sur A, on a :*

(1) *Si X est maximal alors X* est complet.*

(2) *Si X* est complet et X fini alors X est maximal.*

On en déduit :

PROPOSITION 1.2.1 : *Si X est un code maximal fini, pour chaque a de A il existe un entier n tel que a^n appartient à X.*

1.3. Degré d'un code fini

Nous appelons mot infini périodique une application périodique m de $\mathbb{Z}$ dans A. On note :

$$[i, j]\ m = (i)\,m\,(i+1)\ m \ldots (j)\ m.$$

Une interprétation de m par un code X est une application strictement croissante μ de $\mathbb{Z}$ dans $\mathbb{Z}$ telle que : pour tout i de $\mathbb{Z}$ $[i\mu, (i+1)\mu - 1]m$ appartient à X. Si deux interprétations μ_1 et μ_2 de m vérifient $i\mu_1 = j\mu_2$ pour un couple (i, j) de $\mathbb{Z} \times \mathbb{Z}$, on dit qu'elles ont un point commun, sinon qu'elles sont disjointes.

PROPOSITION : *Soit X un code fini; deux interprétations sont disjointes ou confondues.*

LEMME : *L'application ∂ de $\mathbb{Z}$ dans X définie ci-dessous est périodique :*

$$i\mathbb{Z} \dashrightarrow i\partial = [i\mu, (i+1)\mu - 1]m.$$

Preuve : A tout i de $\mathbb{Z}$ on associe par μ un j unique tel que $j\mu \leqq i < (j+1)\mu$ et donc un mot de X. Comme X est fini et m de période p il existe au moins un mot x de X et une infinité d'éléments k de $\mathbb{Z}$ tels que le mot de X associé par μ aux entiers $i + kp$ soit x.

Pour la même raison il existe une infinité d'entiers k' dans $\mathbb{Z}$ tels que les entiers $i + k'p$ désignent la même lettre de ce mot x; enfin la périodicité de m permet de choisir un entier i_0 tel que les entiers $i_0 + k'p$ désignent la première lettre de x.

Posons $m_0 = [i_0, i_0 + p - 1]m$. Les entiers k' sont nécessairement en progression arithmétique, sinon il existerait un mot de la forme m_0^q de X^* qui aurait plusieurs décompositions en mots de X.

Preuve de la proposition : Soit μ_1 et μ_2 deux interprétations de m par X ayant un point commun (i, j); appelons p_1 et p_2 les périodes de ∂_1 et ∂_2 associées à μ_1 et μ_2 et p la période de m.

Posons :

$$k = i\mu_1 = j\mu_2 \qquad \text{et} \qquad m_0 = [k,\ k+p-1]m.$$

Enfin soit q_1 et q_2 les entiers définis par :

$$m_0^{q_1} = [i\mu_1,\ (i+p_1)\mu_1 - 1]m,$$
$$m_0^{q_2} = [j\mu_2,\ (j+p_2)\mu_2 - 1]m.$$

Alors si on désigne par q le ppcm de q_1 et q_2, m_0^q a deux décompositions dans X^* définies à partir de μ_1 et μ_2; ceci implique que μ_1 et μ_2 sont confondues.

DÉFINITION : Le rang de m est le nombre maximal d'interprétations de m. Le degré d'un code fini est le minimum de l'ensemble des valeurs des rangs des mots.

Remarque : Nous parlerons par la suite du rang d'un mot m de A^* au lieu de parler du rang du mot infini périodique m_0 associé en posant :

$$\text{pour tout } k \text{ dans } \mathbb{Z} \quad [kp,\ (k+1)p-1]m_0 = m.$$

DÉFINITION : Un code de degré 1 est dit synchronisant; un code de degré supérieur à 1 est dit asynchrone.

PROPOSITION : *Un code fini est maximal si et seulement si son degré est supérieur ou égal à* 1.

preuve : Supposons que le degré de X soit supérieur ou égal à 1. Alors à tout m de A^* on peut associer un mot infini périodique m_0; m_0 a au moins une interprétation μ et on en déduit immédiatement que m vérifie:

$$A^* m A^* \cap X^* \neq \varnothing.$$

Réciproquement, si X est un code fini maximal, X^* est complet; pour tout m de A^* on a donc :

$$\text{pour tout } n \text{ de } N, \qquad A^* m^n A^* \cap X^* \neq \varnothing.$$

Soit S_x l'ensemble des suffixes propres de X et u, v deux mots de A^* tels que $um^n v$ soit dans X^*. Pour n donné on peut ainsi associer à tout entier i $(1 \leqq i \leqq n)$ le mot s_i de S_x défini par :

$$um^{n-i} = x_1 p, \qquad m^i v = s_i x_2,$$

avec x_1, x_2 dans X^* et ps_i dans X.

X étant fini il existe un entier n_1 tel que l'on ait :

$$\text{il existe } i \text{ et } j, \qquad i \leqq n_1,\ j \leqq n_1,\ s_i = s_j.$$

Ceci nous permet de construire une interprétation du mot infini associé à m.

Remarque 1.3.1 : Si un code maximal contient une lettre de l'aphabet il est synchronisant.

PROPOSITION : *Soit X un code fini synchronisant, il existe un mot x de X^** tel que :

$$x A^* x \subset X^*.$$

Preuve : Soit x un mot de rang 1 dans X^*; pour tout m de A^* et pour tout n entier le mot mx^n a au moins une interprétation par X. En raisonnant comme dans la preuve de la proposition précédente, on montre qu'il existe un entier n tel qu'une interprétation de mx^n permet de construire une interprétation de x. Comme x a une unique interprétation il en est de même pour mx^n. On en déduit qu'il existe deux entiers i et j tels que :

$$x^i m x^j \text{ appartient à } X^*.$$

Remarquons que l'entier n est borné par le cardinal de S_x, n_0, et donc pour tout m de A^* on a :

$$x^{n_0+1} m x^{n_0+1} \text{ appartient à } X^*.$$

COROLLAIRE 1.3.2 : *Soit X un code préfixe, il existe un mot x de X^* tel que : $A^* x \subset X^*$.*

Enfin, si X est un code bipréfixe synchronisant, on doit avoir $A^* \subset X^*$ donc $X = A$.

1.4. Décompositions

DÉFINITION : Un code X sur A se décompose sur un code Y sur A s'il vérifie :

$$X^* \subset Y^* \subset A^* \quad \text{(inclusions strictes)}.$$

On peut alors considérer X^* comme un sous-monoïde libre de Y^* dont la base est un code noté X_y. On écrit : $X = X_y \otimes Y$.

PROPOSITION 1.4.1 : [10] : *Si X et Y sont deux codes maximaux finis sur A tels que* $X = X_y \otimes Y$, *alors :*

$$\text{degré}(X) = \text{degré}(X_y) \times \text{degré}(Y).$$

1.5. Série d'un code

Nous désignons par $Z\langle\langle A \rangle\rangle$ l'anneau des séries formelles en variables non commutatives a de A. Si P est une partie de A^*, on note $\mathbf{P}$ la série caractéristique associée à P.

LEMME : *Soit P et Q deux parties de* A^*. *On a* $\mathbf{P}.\mathbf{Q} = \mathbf{PQ}$ *si et seulement si tout mot de PQ s'écrit d'une façon unique comme produit d'un mot de P par un mot de Q.*

LEMME : *Une partie X de* A^* *engendrant un sous-monoïde L est un code si et seulement si elle vérifie :* $\mathbf{L} = 1/1 - \mathbf{X}$.

DÉFINITION : *Une partie X de* A^* *est factorisante s'il existe deux parties P et Q de* A^* *telles que :* $1 - \mathbf{X} = \mathbf{P}(1 - \mathbf{A})\mathbf{Q}$.

Remarque : Si X est un code factorisant, l'égalité ci-dessus est équivalente à : $\mathbf{A}^* = \mathbf{Q}.\mathbf{X}^*.\mathbf{P}$ qui signifie que tout mot de A^* s'écrit de manière unique sous la forme $q.x.p$ avec q dans Q, p dans P et x dans X^*.

PROPOSITION 1.5.1 : *Une partie factorisante est un code.*

Preuve : Supposons que l'on ait $1 - \mathbf{X} = \mathbf{P}(1 - \mathbf{A})\mathbf{Q}$.

On en déduit l'égalité formelle :

$$1 + \mathbf{A} + \mathbf{A}.\mathbf{A} + \mathbf{A}.\mathbf{A}.\mathbf{A}.\mathbf{A} + \ldots = \mathbf{Q}(1 + \mathbf{X} + \mathbf{X}.\mathbf{X} + \mathbf{X}.\mathbf{X}.\mathbf{X} + \ldots)\,\mathbf{P}.$$

Mais $1 + \mathbf{A} + \mathbf{A}.\mathbf{A} + \ldots = \mathbf{A}^*$. Les séries $\mathbf{P}$, $\mathbf{Q}$ et $\sum(\mathbf{X})^n$ étant à coefficients positifs et la série $\mathbf{A}^*$ à coefficients égaux à $+1$, la série $\sum(\mathbf{X})^n$ est à coefficients égaux à $+1$ et donc :

$$\sum(\mathbf{X})^n = \mathbf{X}^* \quad \text{ou encore} \quad 1/1 - \mathbf{X} = \mathbf{X}^*,$$

X est donc un code.

PROPOSITION 1.5.2 : *Soit X un code factorisant, si P et Q sont de cardinal fini alors X est maximal.*

Preuve : On montre que X^* est complet.

Remarquons que la preuve de la proposition 1.5.1 est encore valide si on suppose seulement que les séries X, P, Q sont à coefficients positifs. On peut

donc dire :

PROPOSITION : Soit P, Q, R, *des séries formelles de Z « A » à coefficients positifs vérifiant :*

(1) *Les valuations de P et Q sont nulles.*

(2) $1-R=P(1-\mathbf{A})Q$.

Alors P, Q, R sont des séries caractéristiques et R représente un code.

PROPOSITION : *Un code préfixe est factorisant.*

Preuve : désignons par D l'ensemble des mots préfixes des mots d'un code préfixe X ou sans préfixe dans X. On a trivialement : $\mathbf{A}^*=\mathbf{X}^*.\mathbf{D}$.

Remarque 1.5.3 : Si X est un code maximal préfixe (resp. suffixe) fini et P_x (resp. S_x) l'ensemble de ses préfixes (resp. suffixes) il vérifie la relation :

$$1-\mathbf{X}=\mathbf{P}_x.(1-\mathbf{A}) \qquad [\text{resp. } 1-\mathbf{X}=(1-\mathbf{A}).\mathbf{S}_x].$$

Remarque 1.5.4 : Les codes préfixes synchronisants ont de plus la propriété de factoriser A^* de manière unique : si X est préfixe synchronisant, il existe (*cf.* 1.3.2) un mot x de X^* tel que A^*x est inclus dans X^*.

Supposons alors que l'on ait $\mathbf{A}^*=\mathbf{X}^*.\mathbf{P}=\mathbf{Q}'.\mathbf{X}^*.\mathbf{P}'$ et soit q' un mot de Q'; $q'x$ est dans X^* et l'unicité de la décomposition d'un mot dans $Q'.X^*.P'$ implique que $q'=1$ et donc $p'=p$.

Par contre un code synchronisant ne factorise pas nécessairement A^* de manière unique.

Exemple : $A=\{a, b\}$:

$$\mathbf{X}=a+ba+b^2, \qquad 1-\mathbf{X}=(1+b)(1-\mathbf{A}),$$
$$\mathbf{X}=a^2+ba+ba^2+b^2a+B^2, \qquad 1-\mathbf{X}=(1+b)(1-\mathbf{A})(1+a).$$

2. CODES INDÉCOMPOSABLES

Nous construisons une famille de codes indécomposables asynchrones.

2.1. Une famille de codes

X désigne ici un code préfixe maximal fini ayant de plus la propriété : il existe un mot w de A^* tel que w^2 appartient à X et :

$$\{m;\ mw\in X\}\cap\{m;\ wm\in X\}=\{w\}.$$

Posons :

$$G=\{m;\ mw\in w\}, \qquad G_1=G\setminus\{w\},$$
$$D=\{m;\ wm\in X\}, \qquad D_1=D\setminus\{w\}.$$

Remarque 2.1.1 : D est un code préfixe maximal.

PROPOSITION : *Les séries T_d et T_g définies ci-dessous représentent des codes :*

$$\mathbf{T}_d=(\mathbf{X}-1+(\mathbf{G}-1)\,w\,(\mathbf{D}-1))\,(1+w)+1,$$
$$\mathbf{T}_g=(1+w)\,(\mathbf{X}-1+(\mathbf{G}-1)\,w\,(\mathbf{D}-1))+1.$$

Preuve : En développant les expressions ci-dessus on obtient :

$$\mathbf{T}_g=\mathbf{X}-\mathbf{G}_1\,w+\mathbf{G}_1\,w\,\mathbf{D}-w\mathbf{D}+w\,(\mathbf{X}-\mathbf{G}w+\mathbf{G}\,w\,\mathbf{D}).$$

$\mathbf{T}_g$ est la série caractéristique d'un ensemble préfixe donc d'un code;

$$\mathbf{T}_d=\mathbf{X}-w\,\mathbf{D}_1+\mathbf{G}\,w\,\mathbf{D}_1-\mathbf{G}\,w+(\mathbf{X}-w\,\mathbf{D}+\mathbf{G}\,w\,\mathbf{D})\,w.$$

$\mathbf{T}_d$ est donc à coefficients positifs. D'autre part appelons P_x et P_d les ensembles des préfixes propres des codes X et D; on peut écrire (*cf.* 1.5.3.) :

$$\mathbf{X}-1=\mathbf{P}_x(\mathbf{A}-1), \qquad \mathbf{D}-1=\mathbf{P}_d(\mathbf{A}-1).$$

En reportant dans l'expression initiale de $\mathbf{T}_d$ on obtient :

$$\mathbf{T}_d=(\mathbf{P}_x-w\,\mathbf{P}_d+\mathbf{G}\,w\,\mathbf{P}_d)\,(\mathbf{A}-1)\,(1+w)+1,$$

$w\,P_d$ est inclus dans P_x car $w\,D$ est inclu dans X. $\mathbf{T}_d$ représente une partie factorisante donc un code.

2.2. Itérabilité de la construction

Le code T_g est un code préfixe maximal contenant le mot w^4.

Posons :

$$G'=\{m;\ mw^2\in T_g\} \qquad \text{et} \qquad D'=\{m;\ w^2\,m\in T_g\}.$$

PROPOSITION : $\mathbf{G}'$ *et* $\mathbf{D}'$ *s'écrivent :*

$$\mathbf{G}'=(1+w)\,\mathbf{G}_1+w^2, \qquad \mathbf{D}'=(1+w)\,\mathbf{D}_1+w^2$$

et ils vérifient : $G'.\,w^2\cap w^2.\,D'=\{w^4\}$.

Preuve : T_g peut s'écrire : $\mathbf{T}_g=(1+w)\ (\mathbf{X}-\mathbf{G}\,w-w\,\mathbf{D}_1+\mathbf{G}\,w\,\mathbf{D}-w^3)+w.$

Posons $\mathbf{F} = \mathbf{X} - \mathbf{G}w - w\mathbf{D}_1 + \mathbf{G}w\mathbf{D} - w^3$; les seuls mots de F qui se terminent par w^2 sont dans $GwD - w^3$. Supposons qu'un mot, gwd, de $GwD - w^3$ se termine par w^2 :

$$gwd = mw^2.$$

Ceci entraîne que le mot wd de X peut s'écrire $m'w$; on aurait alors m' appartient à G. Or par hypothèse $Gw \cap wD = \{w^2\}$. Il s'ensuit que nécessairement $d = w$; on en déduit l'expression de G'.

On montre de la même façon que les mots de F commençant par w^2 sont dans $w^2 D_1$ et que les mots de wF commençant par w^2 sont dans $w^3 D_1$.

Il reste à montrer que les quatre ensembles ci-dessous sont vides :

$$G_1 w^2 \cap w^2 D_1, \qquad G_1 w^2 \cap w^3 D_1,$$
$$w G_1 w^2 \cap w^2 D_1, \qquad w G_1 w^2 \cap w^3 D_1.$$

Pour le premier : X est préfixe et contient w^2 et $G_1 w$; w^2 ne peut donc être préfixe d'un mot de $G_1 w$ ou avoir un préfixe dans $G_1 w$.

On raisonne identiquement pour les trois autres.

En résumé :

PROPOSITION 2.2.1 : *Si un code préfixe maximal fini X contient un mot w^2 à partir duquel on peut construire le code T_g alors la même construction est applicable au code T_g à partir du mot w^4.*

Exemples :

(1) $\mathbf{X} = a + ba + bb$, $w = b$, $\mathbf{G} = b$, $\mathbf{D} = a + b$:

$$\mathbf{T}_g = (1+b)(1+bba) + b^4, \qquad \mathbf{T}_d = (a+bba)(1+b) + b^4.$$

(2) $\mathbf{X} = a + ba(a + ba + b^2a^2 + b^2ab + b^3) + bb$:

$$w = ba, \qquad \mathbf{G} = ba, \qquad \mathbf{D} = a + ba + b^2a^2 + b^2ab + b^3,$$
$$\mathbf{T}_d = (a + bb + baba(a + b^2a^2 + b^2ab + b^3))(1 + ba) + (ba)^4.$$

Le code ici obtenu est synchronisant et indécomposable en codes préfixes ou suffixes.

(3) $\mathbf{X} = aa + ab + baa + baba + babb + bb$:

$$w = ba, \qquad \mathbf{G} = ba, \qquad \mathbf{D} = a + ba + bb,$$
$$\mathbf{T}_d = (aa + ab + bb + baba(a + bb))(1 + ba) + (ba)^4.$$

T_d est synchronisant, indécomposable en codes préfixes ou suffixes.

2.3. Degré des codes construits

Remarquons que T_d et T_g peuvent s'écrire :

$$\mathbf{T}_d = \mathbf{F}(1+w)+w^4, \qquad \mathbf{T}_g = (1+w)\mathbf{F}+w^4,$$

avec $\mathbf{F} = \mathbf{X} - w\mathbf{D} - \mathbf{G}_1 w + \mathbf{G}w\mathbf{D} - w^3$.

Tout mot de T_d (resp. T_g) distinct de w^4 s'écrit donc de façon unique fw^k (resp. $w^k f$) avec k dans $\{0, 1\}$ et f dans F.

PROPOSITION 2.3.1 : $\mathbf{T}_d$ et $\mathbf{T}_g$ *ont le même degré.*

Preuve : Soit μ_d une interprétation d'un mot infini périodique m par T_d; on lui associe une interprétation μ_g de m par T_g comme suit :

Soit i dans $\mathbb{Z}$ et soit j le plus grand entier inférieur à i vérifiant (s'il existe) :

$$[j\mu_d, (j+1)\mu_d - 1]\, m = fw^k,$$

on pose : $i\mu_g = i\mu_d - k \,.\, |w|$.

Si j n'existe pas on pose : $i\mu_g = i\mu_d$.

Symétriquement on peut associer une interprétation de m par T_d à une interprétation de m par T_g.

m étant infini périodique, si un mot de la forme fw^k apparaît une fois dans une interprétation, alors il apparaît une infinité de fois. On en déduit que l'application définie entre les interprétations de m par T_d et T_g est bijective et donc que ces codes ont le même degré.

Remarque : T_d et T_g n'ont pas forcément le même degré que X, comme le montre l'exemple suivant :

$\mathbf{X} = (a+ba+bb)^2$ X est préfixe de degré 2, composé du code $\mathbf{X}_y = (t+u+v)^2$ sur l'alphabet $\{t, u, v\}$ et du code $\mathbf{Y} = a+ba+bb$.

Choisissons $w = ba$; alors :

$$\mathbf{G} = a+b+ba+bb, \qquad \mathbf{D} = a+ba+bb.$$

On trouve :

$$\mathbf{T}_g = (1+ba)(1+a+bb+aba+bba+baba+bbba)(a+ba+bb-1)+1,$$

cette expression peut s'écrire comme un polynome sur $\{t, u, v\}$:

$$\mathbf{T}_y = (1+u)(1+t+v+tu+vt+uu+vu)(t+u+v-1)+1.$$

T_y est un code préfixe et $T_g = T_y \otimes Y$ de plus T_y est synchronisant, donc aussi T_g (*cf.* 1.4.1).

2.4. Décomposabilité sur un code synchronisant

Imposons maintenant que w^2 ne soit pas égal à une puissance d'une lettre de A. Les codes T_d et T_g contiennent pour tout a de A le mot a^{n_a} de X.

PROPOSITION : *Si T_g est un code asynchrone et si les entiers n_a sont premiers, T_d et T_g sont indécomposables sur un code synchronisant.*

Preuve : Supposons que T_g se décompose sur un code synchronisant Y :

$$T_g = Z \otimes Y, \qquad \text{degré}(Y) = 1,$$

ceci entraîne degré $(T_g) \neq 1$ car T_g est maximal fini.

a^{n_a} étant dans T_g il est dans Y^*; on ne peut avoir a^{n_a} dans Y car alors le code Z sur l'alphabet Y contiendrait une lettre, a^{n_a}, et serait donc synchronisant (*cf.* 1.3.1); il s'ensuit que a est dans Y pour tout a de A, c'est-à-dire $Y = A$. La preuve est identique pour T_d (*cf.* 2.3.1).

2.5. Codes généraux indécomposables

Nous choisissons pour X un code bipréfixe maximal fini de degré premier d; pour tout a de A on a alors a^d appartient à X.

LEMME : *Si X contient un mot de la forme w^2, il existe pour tout a de A un entier p_a (resp. q_a) tel que a^{p_a} appartient au code D (resp. a^{q_a} appartient à G) et $p_a < d$ (resp. $q_a < d$).*

La preuve de ces résultats est une conséquence de la bipréfixité et de la maximalité de X.

Supposons maintenant que X vérifie la propriété : $Gw \cap wD = \{w^2\}$ on peut construire le code préfixe T_g :

$$\mathbf{T}_g = (1+w)(\mathbf{X} - 1 + (\mathbf{G}-1)\, w\, (\mathbf{D}-1)) + 1$$

et à partir de w^4 et T_g le code I_d :

$$\mathbf{I}_d = (\mathbf{T}_g - 1 + (\mathbf{G}'-1)\, w^2\, (\mathbf{D}'-1))\,(1+w^2) + 1,$$

avec :

$$\mathbf{G}' - 1 = (1+w)(\mathbf{G}-1), \qquad \mathbf{D}' - 1 = (1+w)(\mathbf{D}-1) \quad (cf.\ 2.2.1.),$$

ce qui donne :

$$\mathbf{I}_d = (1+w)(\mathbf{X} - 1 + (\mathbf{G}-1)(w+w^2+w^3)(\mathbf{D}-1))\,(1+w^2) + 1,$$

après développement et réductions on obtient :

$$\mathbf{I}_d = (1+w)(\mathbf{X} - \mathbf{G}w - w\mathbf{D}_1 + \mathbf{G}_1 w\mathbf{D}_1 + \mathbf{G}_1 w^2\mathbf{D}_1 + \mathbf{G}w^3\mathbf{D} - w^5)(1+w^2) + w^8.$$

PROPOSITION : *Si X est de degré premier différent de 2, le code I_d est indécomposable sur un code préfixe ou suffixe.*

Preuve : Supposons que I_d se décompose sur un code préfixe Y; I_d contient les mots w^2 et $a^d w^2$ pour tout a de A. Donc w^2 est un mot de Y^*; comme $Gw^3D \setminus \{w^5\}$ est une partie de I_d et Y est préfixe, on déduit :

$D \setminus \{w\}$ est une partie de Y^* et donc a^{p_a} appartient à Y^* pour tout a de A.

En résumé : Pour tout a de A, a^d et a^{p_a} sont dans Y^* et $p_a < d$. Ceci entraîne que a est dans Y et donc $Y = A$.

On raisonne symétriquement si Y est supposé suffixe.

PROPOSITION : *I_d est un code asynchrone.*

Preuve : Les conditions imposées à X permettent de construire aussi le code I_g à partir de T_d :

$$\mathbf{I}_g = (1+w^2)(\mathbf{T}_g - 1 + (\mathbf{G}' - 1)w^2(\mathbf{D}' - 1)) + 1,$$

qui s'écrit encore :

$$\mathbf{I}_g = (1+w^2)(1+w)(\mathbf{X} - 1 + (\mathbf{G} - 1)(w + w^2 + w^3)(\mathbf{D} - 1)) + 1.$$

X et G étant des codes suffixes maximaux, on peut les factoriser (*cf.* 1.5.3.) :

$$\mathbf{X} - 1 = (\mathbf{A} - 1)\mathbf{S}_x, \qquad \mathbf{G} - 1 = (\mathbf{A} - 1)\mathbf{S}_g.$$

On peut donc factoriser I_g sous la forme :

$$\mathbf{I}_g = (1+w^2)(1+w)(\mathbf{A} - 1)(\mathbf{S}_x + \mathbf{S}_g(w + w^2 + w^3)(\mathbf{D} - 1)) + 1$$

et aussi :

$$\mathbf{I}_g = (1+w^2)(1+w)(\mathbf{P}_x + (\mathbf{G} - 1)(w + w^2 + w^3)\mathbf{P}_d)(\mathbf{A} - 1) + 1.$$

I_g admet ainsi plusieurs factorisations, il est asynchrone (*cf.* 1.5.3). Nous savons d'autre part que I_g et I_d ont le même degré.

Exemple : Soit X le code bipréfixe de degré 3 sur $\{a, b\}$ défini par :

$$\mathbf{X} = aaa + aaba + aabb + ab + baa + baba + babb + bba + bbb,$$

choisisons $w = ba$.

$$\mathbf{G} = aa + ba + b, \qquad \mathbf{D} = a + ba + bb,$$
$$\mathbf{I}_d = (1 + ba)\,\mathbf{F}(1 + (ba)^2) + (ba)^8,$$

avec :

$$\mathbf{F} = a^3 + a^2 b^2 + ab + b^3 + (a^2 + b)(ba + baba)(a + b^2) + (a^2 + ba + b)(ba)^3 (a + ba + b^2) - (ba)^5.$$

Remarque [10] : Pour tout entier d premier il existe au moins un code bipréfixe vérifiant nos hypothéses sur $A = \{a, b\}$:

$$\mathbf{X} = \mathbf{A}^d + (\mathbf{A} - 1)\, a^{n-1}\, ba^n\, (\mathbf{A} - 1),$$

avec $d = 2n + 1$ on prend $w = a^n b$.

REMERCIEMENTS

Je voudrais enfin remercier le rapporteur anonyme pour ses nombreuses remarques et suggestions qui m'ont permis d'améliorer cet article.

BIBLIOGRAPHIE

1. J. M. Boë, *Une famille remarquable de codes indécomposables*, Automata, Languages and Programming, Ausiello and Böhm, Springer-Verlag, 1978, p. 105-112.
2. J. M. Boë, *Représentations des monoïdes. Application à la théorie des codes*, Thèse 3e cycle, 1976, U.S.T.L., Montpellier.
3. J. Boyat, *Sur la construction des codes à longueur variable au moyen des représentations de leur monoïde syntaxique*, Thése 3e cycle, 1977, U.S.T.L., Montpellier.
4. Y. Césari, *Sur l'application du théorème de Suschkevitch à l'étude des codes rationnels complets*, Automata, Languages and Programming, Loeckx, 1974, p. 342-350.
5. Y. Césati, *Sur un algorithme donnant les codes bipréfixes finis*, Math. System. Theory, 1972, p. 221-225.
6. S. Eilenberg, *Automata, Langages and Machines*, vol. A, Academic Press, New York, 1974.
7. G. Lallement, *Semigroups and Combinatorial Applications*, John Wiley and Sons, New York, 1979.

8. M. NIVAT, *Éléments de la théorie générale des codes*, Automata Theory, Academic Press, New York, 1966, p. 278-294.
9. D. PERRIN, *Codes bipréfixes et groupes de permutation*, Thèse, Paris, 1975.
10. D. PERRIN, *Codes asynchrones*, Bull. Soc. Math. Fr., vol. 105, 1977, p. 385-404.
11. J. F. PERROT, *The Theory of Variable Length Codes*, Theoretical Computer Science, Springer-Verlag, 1976, p. 24-27.
12. M. P. SCHÜTZENBERGER, *Une théorie algébrique du codage*, Séminaire Dubreil-Pisot, 1955-1956, 15.
13. M. P. SCHÜTZENBERGER, *On a Special Class of Recurrent Events*, Ann. Math. Stat., vol. 32, 1961, p. 1201-1213.
14. M. P. SCHÜTZENBERGER, *Sur certains sous-monoïdes libres*, Bull. Soc. Math. Fr., vol. 93, 1965, p. 209-223.
15. M. P. SCHÜTZENBERGER, *Sur le produit de concaténation non ambigu*, Semigroup forum, vol. 13, 1976, p. 47-75.
16. C. E. SHANNON, *A Mathematical Theory of Communication*, Bull. Syst. Techn. J., vol. 27, 1948, p. 379-423.