Parties de $\mathbb{N}$ reconnaissables
Théorème de Cobham

Ivan Boyer

12 février 2007

Table des matières

1 Définitions et notations

Soit A un alphabet fini et $k = |A|$. Grâce à une bijection $f : A \to [\![0,k-1]\!]$ on peut représenter un entier n écrit en base k à l'aide des lettres de A. On sous-entend par représentation, une représentation « propre », c'est-à-dire sans des $f^{-1}(0)$ à gauche des mots. Par exemple, nous écrivons communément les entiers en base 10 à l'aide de l'alphabet $A = \{0, \ldots, 9\}$ par des représentations propres : on n'écrit pas 007 mais simplement 7. Avec ces conventions, on introduit la notion principale :

Définition 1.1 (Ensemble k–rationnels)**.** *Soit $k > 1$.*

Une partie $P \subset \mathbb{N}$ est dite k–rationnelle s'il existe un alphabet A de taille k et un automate qui reconnaisse les entiers de P écrits en base k grâce aux lettres de A.

Pour $k = 1$ on prolonge cette définition en représentant un entier n sur un alphabet unaire, $A = \{a\}$, par le mot a^n. On dira simplement que P est rationnelle, plutôt que 1–rationnelle.

Notons que l'alphabet importe peu pour représenter des entiers en base k. En effet, pour une partie $P \subset \mathbb{N}$ représentée en base k dans deux alphabets A et B on a un morphisme naturel entre les représentations sur A et sur B. Ainsi, dans la suite, on ne mentionnera plus les alphabets. De plus, on assimilera 0 et la lettre représentant 0 dans une base quelconque. Enfin, on utilisera la notation :

Notation 1.2. *Soit P une partie de $\mathbb{N}$. On notera $(P)_k$ l'ensemble de ses représentations en base k.*

Donnons de plus une définition clé pour le théorème de Cobham :

Définition 1.3. *Deux entiers k et l sont dit multiplicativement indépendants s'il n'existe pas d'entiers $n, p > 0$ tels que $k^n = l^p$.*

Remarquons que c'est une notion assez « faible » puisque pour être multiplicativement dépendants il ne suffit pas que les facteurs premiers de l'un soient exactement les facteurs premiers de l'autre. Par exemple, 10 et 20 sont multiplicativement indépendants.

Théorème (Cobham)**.** *Soient k et l multiplicativement indépendants. Une partie $P \subset \mathbb{N}$ est k– et l–rationnelle si et seulement si elle est rationnelle.*

Notons que c'est le sens direct qui constitue généralement ce que l'on entend par théorème de Cobham. Le sens indirect est beaucoup plus facile et sera montré au paragraphe 2.

Exemple 1.1. L'ensemble $\{3, 5, 9, 17, 33, \ldots\}$ n'est certainement pas rationnel car le lemme de l'étoile impose d'avoir un sous ensemble en progression arithmétique. Comme il est 2–rationnel (c'est 10*1), il n'est donc pas 10–rationnel.

Rappelons par ailleurs une autre définition classique sur les parties de $\mathbb{N}$:

Définition 1.4. *Une partie $P \subset \mathbb{N}$ est dite ultimement périodique s'il existe des entiers N et $k > 0$ tels que*

$$\forall n \geqslant N,\ n \in P \Leftrightarrow n + k \in P$$

Notons que toute partie finie est ultimement périodique et qu'une modification d'un nombre fini de nombres n'altère pas le caractère ultimement périodique.

Enfin, la démonstration du théorème de Cobham utilise le résultat suivant de densité :

Théorème 1.1. *Soient k, l deux entiers multiplicativement indépendants. Alors, $\left\{\frac{k^p}{l^q},\ (p,q) \in \mathbb{N}^2\right\}$ est dense dans $\mathbb{R}^+$.*

Pour montrer ce théorème, nous montrons auparavant le lemme suivant :

Lemme 1.2. *Soit α un irrationnel. Alors, l'ensemble $\{p - q\alpha,\ p, q \in \mathbb{N}\}$ est dense dans $\mathbb{R}$.*

<u>Preuve</u> :
► Soit $\alpha_n = \frac{p_n}{q_n}$ une suite de rationnels, sous forme irréductible, tendant vers α. Comme α est irrationnel, on peut choisir $(p_n)_{n\in\mathbb{N}}$ et $(q_n)_{n\in\mathbb{N}}$ strictement croissantes. On peut aussi supposer que $0 < \alpha_n - \alpha < \frac{1}{q_n}$. Ainsi, $\beta_n = p_n - \alpha q_n \in [0,1]$. Par Bolzano-Wierstrass, quitte à extraire une sous-suite, on supposera que β_n converge. Dès lors, en considérant $\beta_{n+1} - \beta_n$, on a trouvé $p_n, q_n > 0$ tels que $p_n - q_n\alpha$ converge vers 0, par valeurs positives. Soit $x \in \mathbb{R}^+$ un réel : on va l'approcher à $\epsilon > 0$. Il existe un rang N tel que $0 < p_N - q_N\alpha < \epsilon$ et on peut alors trouver un entier $k_N > 0$ tel que $0 < k_N p_N - k_N q_N \alpha - x < \epsilon$.

Si le réel x à approcher est négatif, on prend $q_0 > 0$ tel que $q_0\alpha > x$ et on approche comme ci-dessus $q_0\alpha - x$. ◄

Notons que quitte à choisir un plus petit ϵ, on peut prendre p_n et q_n aussi grands que l'on veut. Ceci nous sera utile plus tard. Montrons maintenant le théorème :

<u>Preuve</u> (du théorème) :
► On a $\frac{\log k}{\log l} = \frac{q}{p} \Leftrightarrow k^p = l^q$. Ainsi, si k et l sont multiplicativement indépendants, alors $\frac{\log k}{\log l}$ est irrationnel. Dès lors, pour $x > 0$, on peut trouver, grâce au lemme précédent, deux entiers p_n et q_n tels que :

$$p_n \log k - q_n \log l - \log x$$

soit aussi petit que l'on veut. ◄

2 Résultats préliminaires

Avant de s'attaquer au théorème de Cobham voyons quelques résultats plus élémentaires sur les parties de $\mathbb{N}$ k–rationnelles.

La première est une conséquence assez intuitive du lemme de l'étoile :

Proposition 2.1. *Une partie $P \subset \mathbb{N}$ est rationnelle si et seulement si elle est ultimement périodique.*

<u>Preuve</u> :

▶ Le sens indirect s'obtient par la construction évidente de l'automate.

Le sens direct est une conséquence du lemme de l'étoile : soit k le plus petit entier[1], tel qu'il existe $n \geqslant 0$ vérifiant $a^n \left(a^k\right)^* \subset P$. Soit X l'ensemble des entiers $n < k$ vérifiant l'inclusion précédente. Alors,

$$P \setminus \bigcup_{n \in X} a^n \left(a^k\right)^*$$

est rationnel. Si jamais il n'est pas fini, on peut de nouveau appliquer le lemme de l'étoile qui fournit un entier $k' \notin k\mathbb{N}$ tel que $a^n \left(a^{k'}\right)^* \subset P$ et le pgcd de k et k' contredit la définition de k ◀

Montrons maintenant le sens indirect du théorème de Cobham :

Proposition 2.2. *Si $P \subset \mathbb{N}$ est rationnelle, elle est k–rationnelle pour $k \geqslant 1$.*

<u>Preuve</u> :

▶ Pour faciliter la rédaction, on se placera pour la k–rationnalité sur l'alphabet $[\![0,k-1]\!]$. D'après la proposition précédente, il suffit de montrer le résultat pour les ensembles d'entiers qui ont un reste égal à s modulo p. On construit un automate ayant pour ensemble d'états $[\![0,q-1]\!]$. L'état initial est 0 et l'état final s. On met les transitions $q \xrightarrow{a} r$ pour $r \equiv qk + a[p]$. ◀

Exemple 2.1.

Pour illustrer cette proposition, voici ci-contre l'automate construit comme dans la preuve pour la reconnaissance des multiples de 3 en base 10. On remarquera que l'on trouve l'automate basé sur le critère usuel de divisibilité par 3.

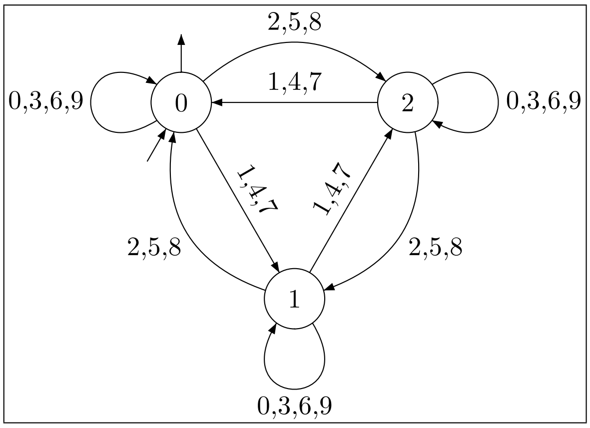

FIG. 1 – Multiples de 3 en base 10.

[1] k existe dès que P est infinie ; le cas P finie est trivial.

Voici enfin une proposition qui illustre l'enjeu des entiers multiplicativement indépendants :

Proposition 2.3. *Une partie $P \subset \mathbb{N}$ est k–rationnelle si et seulement si elle est k^p–rationnelle, pour $p > 0$ arbitraire.*

<u>*Preuve*</u> :

► Le passage de k^p à k est facile, en ajoutant pour chaque transition des états et transitions intermédiaires pour écrire les « chiffres » de la base k^p en base k.

Supposons que P soit k–rationnelle, reconnue par un automate $\mathcal{A}$. On construit un nouvel automate ayant les mêmes états que $\mathcal{A}$. Pour chaque suite de p transitions entre deux états q et q', on met une transition entre q et q' étiquetée par le chiffre de la base k^p représenté par les p transitions. ◄

Ainsi si k et l sont multiplicativement dépendants alors une partie de $\mathbb{N}$ est k–rationnelle si et seulement si elle est l–rationnelle.

3 Lemmes intermédiaires

La démonstration du théorème de Cobham présentée ci-dessous[2] est assez longue, divisée en plusieurs lemmes. Le premier utilise la notion de densité à droite :

Définition 3.1. *Soit A un alphabet et L un langage sur A. L est dit dense à droite si tout mot $u \in A^*$ apparaît comme préfixe d'un mot de L.*

Lemme 3.1. *Soit $P \subset \mathbb{N}$ k–rationnelle et infinie. Alors, pour tout entier l multiplicativement indépendant, l'ensemble $0^*(P)_l$ est dense à droite.*[3]

<u>*Preuve*</u> :

► On peut appliquer le lemme de l'étoile à $(P)_k$ pour trouver $u, v, w \in A^*$ avec $v \neq \varepsilon$ tels que $uv^*w \subset (P)_k$. Soit x une représentation[4] en base k d'un entier arbitraire. Il faut montrer que x apparaît comme préfixe d'un mot de $0^*(P)_l$.

Notons n_x, n_u, n_v les entiers représentés respectivement par x, u et v. Notons $g = |v|$ et $h = |w|$. D'après le théorème 1.1 appliqué aux entiers k^g et l, on peut trouver p et q tels que :

$$\frac{n_x + \frac{1}{4}}{k^h} < \left(n_u + \frac{n_v}{k^g - 1}\right)\frac{k^{gp}}{l^q} < \frac{n_x + \frac{1}{2}}{k^h}$$

[2] due à G. Hansel

[3] Les 0 ne sont là que pour les représentations impropres.

[4] Pas forcément la représentation propre.

Notons, comme on l'a déjà dit dans le lemme 1.2, que l'on peut prendre p et q aussi grands que l'on veut. On peut donc choisir q suffisamment grand pour que :

$$-\frac{1}{4} < \frac{n_w - \frac{k^h n_v}{k^g-1}}{l^q} < \frac{1}{2}$$

En remarquant que la quantité :

$$\begin{aligned}\left(n_u + \frac{n_v}{k^g-1}\right)k^{gp+h} + n_w - \frac{k^h n_v}{k^g-1} &= n_u k^{pg+h} + n_v \frac{k^{gp}-1}{k^g-1}k^h + n_w \\ &= n_u k^{pg+h} + \left(k^h \sum_{i=0}^{p-1} n_u k^{gi}\right) + n_w\end{aligned}$$

n'est autre que l'entier représenté par $uv^p w$ (noté $n_{uv^p w}$), on obtient, en sommant les deux inégalités :

$$n_x l^q < n_{uv^p w} < (n_x+1)l^q$$

Ceci montre l'existence de $j \in [\![1, l^q - 1]\!]$ tels que $n_x l^q + j = n_{uv^p w} \in P$. Ainsi x est bien le préfixe d'un mot de $(P)_l$, à savoir $uv^p w$. ◀

Pour le lemme suivant, on introduit une nouvelle notion :

Définition 3.2. *On dira qu'une partie non vide $P \subset \mathbb{N}$ est presque-périodique de période $d > 0$ si pour tout $x \in P$ on peut trouver $y \in P$ tels que $x < y \leqslant x + d$.*

Remarquons que si P est *presque-périodique* de période $d > 0$ alors, tout intervalle d'entiers supérieurs à $\min P$ de longueur d contient au moins un élément de P.

Lemme 3.2. *Soit $P \subset \mathbb{N}$ k–rationnelle. P est presque-périodique si et seulement si $0^*(P)_k$ est dense à droite.*

<u>Preuve</u> :

▶ Le sens direct n'utilise pas l'hypothèse : soit $n \in \mathbb{N}$ et u sa représentation propre en base k. Soit d la *presque-période* de P et p son nombre de lettres une fois écrite en base k. Il existe $p' \geqslant p$ tel que $k^{p'} \geqslant \min P$: on peut trouver $m \in P$ tel que $nk^{p'} < m < (n+1)k^{p'}$.

Le sens indirect est aussi rapide : pour tout n, il existe deux entiers p et $t < k^p$ tels que $nk^p + t \in P$. On note q le nombre d'états d'un automate reconnaissant $(P)_k$. Une fois que l'automate a lu u (représentant n), l'existence de t montre que l'on peut atteindre un état final. Comme on peut le faire sans repasser deux fois par le même état, on peut prendre $p \leqslant q$ et donc P est *presque-périodique*, de période k^q ◀

Ceci montre déjà une version faible du théorème de Cobham, à savoir :

Proposition 3.3. *Soit $P \subset \mathbb{N}$ infini. Si P est k–rationnelle et l–rationnelle, pour k et l multiplicativement indépendants, alors P est presque périodique.*

Il nous faut donc étudier un peu plus les parties de $\mathbb{N}$ *presque-périodiques.* Voici un lemme, surtout technique, sur ces parties *presque-périodiques* :

Lemme 3.4. *Soit $P \subset \mathbb{N}$ presque-périodique de période d. Pour tout entier K, L et h, et pour tout réel $\eta > 0$ vérifiant $K < L < K + \eta$, il existe $x \in P$ et $y \in \mathbb{N}$ tels que :*

$$yL \leqslant xK + h \leqslant yL + \eta d$$

Preuve :

▶ Soit $r = \min\{p \in \mathbb{N},\ pL > pK + h\}$. Soit $i \in [\![1,r-1]\!]$. On a d'une part $(r-i)L < (r-i)K + h$ et d'autre part :

$$(r-i)K + h = rK + h - iK < rL - iK < rL - iL + i\eta = (r-i)L + \eta i$$

Remarquons que pour $i \geqslant r$ ces inégalités restent vraies. Soit j un entier vérifiant $jK + r - d \geqslant 0$ et $jL + r - d \geqslant \min P$. En ajoutant jKL aux différentes inégalités, on a finalement pour tout $i \in [\![1,d]\!]$:

$$(jK + r - i)L < (jL + r - i)K + h < (jK + r - i)L + \eta d$$

Comme P a d comme *presque-période* alors, il existe un $i \in [\![1,d]\!]$ tel que $x := jL + r - i \in P$. En notant enfin $y := jK + r - i$, on a le résultat annoncé. ◀

Introduisons maintenant une notion clé dans cette preuve du théorème de Cobham :

Définition 3.3 (Facteurs récurrents)**.** *Soit x un mot infini et w un facteur de x c'est-à-dire $x = awb$. S'il existe une infinité de tels couples (a, b) vérifiant cette relation, le facteur w est dit récurrent.*
De plus, on dira qu'un mot $x_0x_1\ldots$ est ultimement périodique[5] *si :*

$$\exists N, p \in \mathbb{N}, \forall n \geqslant N,\ x_n = x_{n+p}$$

Le lien entre les deux notions de cette définition est énoncé dans le lemme suivant :

Lemme 3.5. *Un mot x est ultimement périodique si et seulement si on peut trouver un entier $n > 0$ tel que x ait au plus n facteurs récurrents de longueur n.*

Preuve :

▶ Le sens direct est assez immédiat puisque pour un un mot ultimement périodique de période p, le nombre de facteurs récurrents de longueur p est bien au plus p.

[5] On utilise le même vocabulaire car il n'y a pas d'ambiguïté possible.

Soit $r(n)$ le nombre de facteurs récurrents de longueur n. Chacun de ces facteurs apparaît une infinité de fois dans x et comme l'alphabet est fini, il y a encore une infinité de positions où le facteur est suivi d'une même lettre. Ainsi, $r(n) \leqslant r(n+1)$. Par hypothèse, il existe un plus petit entier $n_0 \geqslant 1$ tel que $r(n_0) = n_0$. Si $n_0 = 1$, le mot est constant à partir d'un certain rang et donc ultimement périodique. Sinon $n_0 \geqslant 2$. On a alors $r(n_0 - 1) = n_0$. Ainsi, au bout d'un moment, chaque facteur de longueur $n_0 - 1$ est suivi d'une unique lettre ne dépendant que du facteur : pour n assez grand et $w = x_n \dots x_{n+n_0-2}$ facteur récurrent, x_{n+n_0-1} est déterminé uniquement par w et $w' = x_{n+1} \dots x_{n+n_0-1}$ est encore un facteur récurrent qui détermine parfaitement x_{n+n_0}. On conclut ainsi que x est ultimement périodique.

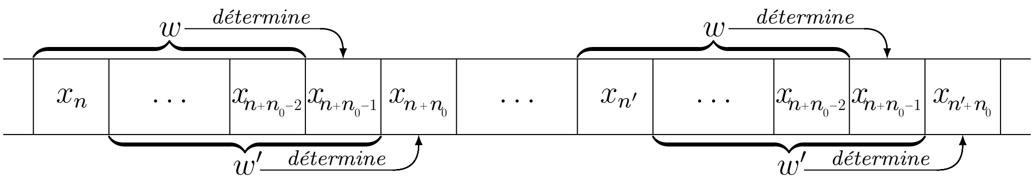

FIG. 2 – x est ultimement périodique de période $n' - n$

◀

4 Théorème de Cobham

Nous avons maintenant les outils nécessaires pour pouvoir prouver le sens direct le théorème de Cobham

Théorème (Cobham)**.** *Une partie $P \subset \mathbb{N}$ k– et l–rationnelle est rationnelle dès que k et l sont multiplicativement indépendants.*

Preuve :

▶ L'ensemble $E_{tj} = \{y,\ yk^j + t \in P\}$ est k–rationnel et aussi l–rationnel. En effet, la constante t n'affecte qu'un nombre fini d'états sur la droite et la division par k^j s'effectue classiquement en rajoutant une sortie sur les transitions de l'automate.

Ainsi, pour tout entier n, représenté en base k par un mot u, l'ensemble des entiers ayant une écriture en base k dans $(P)_k u^{-1} = \{v,\ vu \in (P)_k\}$ est l–rationnel, puisqu'il s'agit de $E_{n|u|}$

Soit $\sim$ la relation d'équivalence sur $\mathbb{N}$ définie par $x \sim y$ si et seulement si $(x)_k^{-1}(P)_k = (y)_k^{-1}(P)_k$. Notons que l'on a la propriété

$$x \sim y \Rightarrow xk^j + t \sim yk^j + t \tag{1}$$

pour tout j et tout t tels que $t < k^j$. Remarquons aussi que :

$$x \sim 0 \Leftrightarrow (x)_k^{-1}(P)_k = (P)_k \Leftrightarrow x \in P \tag{2}$$

On vérifie facilement que la classe d'un entier x a pour représentation en base k :

$$([x]_\sim)_k = \left(\bigcap_{u\in(x)_k^{-1}(P)_k} (P)_k u^{-1}\right) \bigcap \left(\bigcap_{u\notin(x)_k^{-1}(P)_k} {}^c\big((P)_k u^{-1}\big)\right)$$

Comme le nombre de quotients à droite est fini, les classes d'équivalence de $\sim$ sont l–rationnelles.

Dès lors, pour chacune de ces classes écrite en base l, le nombre de quotients à droite est fini et on peut donc raffiner la relation $\sim$ en $\approx$ ayant toujours un nombre fini de classes, que l'on note c, telle que :

$$x \approx y \Rightarrow xl^j + t \approx yl^j + t \tag{3}$$

pour tout j et tout t tels que $t < l^j$.

On numérote alors les classes modulo $\sim$ et on note $num([k]_\sim)$ le numéro de la classe de k modulo $\sim$. On considère le mot infini R défini par $R_n = num([n]_\sim)$.

Soit w un facteur récurrent de longueur 2 et $P_w = \{n,\ R_nR_{n+1} = w\}$. P_w est k– et l–rationnel (en utilisant un automate produit où les transitions sont celles de la classe w_0, doublées de celle de la classe w_1). Ainsi, d'après le lemme 3.3, P_w est *presque-périodique* de période d_w : chaque occurrence de w est suivi d'une autre dans un intervalle d'au plus d_w. Comme le nombre de facteurs récurrents de longueur 2 est fini, on peut définir $d = \max d_w$. Soit alors ϵ un réel vérifiant $0 < \epsilon < 1$ ainsi que $c\frac{\epsilon}{1-\epsilon} < \frac{1}{2}$. Grâce théorème 1.1 on trouve deux entiers p et q tels que :

$$1 < \frac{l^q}{k^p} < 1 + \frac{\epsilon}{d}$$

Soit maintenant $K = k^p$, $L = l^q$ et $m = \lfloor K(1-\epsilon)\rfloor$. Nous avons besoin d'un dernier lemme :

Lemme 4.1. *Soit w un facteur récurrent de longueur m de R. Il existe un entier y tel que $R_{yL}\dots R_{(y+1)L-1} = swt$ et $|s| \leqslant \epsilon K$*

<u>*Preuve*</u> :

▶ Il existe une infinité d'entiers x tel que $R_{xK}\dots R_{(x+2)K-1}$ ait w pour facteur. On peut de plus imposer que w commence toujours à la même position. Notons que la relation $\sim$ vérifie la propriété 1 et donc que le sous mot $R_{xK}\dots R_{(x+2)K-1}$ est entièrement défini par le facteur R_xR_{x+1} et par K. Parmi tous les facteurs R_xR_{x+1}, il y en a nécessairement un qui est récurrent : on est en mesure de construire une suite strictement croissante $(x_n)_{n\in\mathbb{N}}$ vérifiant :

$$\begin{cases} \forall n,\ x_{n+1} - x_n \leqslant d \\ R_{x_nK}\dots R_{(x_n+2)K-1} = w'_n w w''_n \end{cases}$$

avec $|w'_n| = h$ constant. On peut maintenant utiliser le lemme 3.4 avec h ainsi défini, $\eta = K\frac{\epsilon}{d}$ et l'ensemble P *presque-périodique*[6]. On vérifie que l'on a bien $K < L < K + \eta$. On a donc l'existence d'un entier y tel que $yL < xK + h < yL + \eta d < (y+1)L$. Ainsi, on a :

$$R_{yL} \dots R_{(y+1)L-1} = swt$$

avec $|s| < yL + \eta d - yL$. La figure suivante aide à se retrouver parmi tous ces indices.

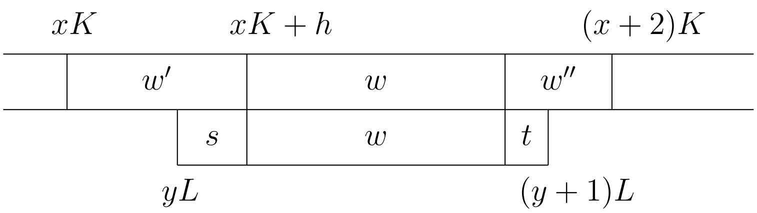

Fig. 3 – Localisation de w. ◀

Les facteurs de la forme $R_{yL} \dots R_{(y+1)L-1}$ seraient parfaitement définis par R_y et L si $\sim$ possédait la propriété 3. Or $\approx$ possède précisément cette propriété. Ainsi, il y a au plus c possibilités de facteurs $R_{yL} \dots R_{(y+1)L-1}$, correspondants à la classe de R_y modulo $\approx$. Comme L ne dépend pas de w, un majorant du nombre de facteurs récurrents de R de longueur m est donné par un facteur de la forme $R_{yL} \dots R_{(y+1)L-1}$ ainsi que par le point de départ de w dans ce facteur. Ce point de départ étant fourni par $|s|$, on a donc au plus

$$\epsilon K \times c \leqslant \frac{1}{2}K(1-\epsilon) \leqslant \frac{1}{2}(m+1) \leqslant m$$

facteurs récurrents de R de longueur m. Ainsi, d'après le lemme 3.5, le mot R est ultimement périodique. Or, vu la propriété 2, P est précisément la classe de 0 modulo $\sim$ c'est-à-dire $P = \{n,\ R_n = num([0]_\sim)\}$. Ainsi, il est clair que P est elle-même ultimement périodique, ce qui veut dire, d'après la proposition 2.1 que P est rationnelle. ◀

5 Autres résultats de reconnaissabilité

5.1 Reconnaissance par substitution

La notion de partie k–rationnelle peut être définie d'une autre façon à l'aide de morphismes substitutifs.

[6]D'après la proposition 3.3

Définition 5.1. *Une (k–)substitution est un morphisme $\alpha : A \to A^k$ tel qu'au moins une lettre $a \in A$ soit la première lettre de $\alpha(a)$. Un point fixe de α est un mot infini, w, tel que $\alpha(w) = w$.*

Notons que pour chaque lettre $a \in A$ qui est la première lettre de $\alpha(a)$ il existe un unique point fixe w commençant par a, à savoir $\alpha^\omega(a)$. Nous avons la proposition suivante :

Proposition 5.1. *Une partie $P \subset \mathbb{N}$ est k–rationnelle si, et seulement si, il existe un alphabet fini A et une k–substitution ayant un point fixe w tel que :*

$$P = \bigcup_{a \in A_P} \{n \geqslant 0,\ w_n = a\}$$

où $A_P \subset A$ est la partie de A qui "détermine" P

<u>Preuve</u> :

▶ Supposons P k–rationnelle sur l'alphabet $[\![0,k-1]\!]$, reconnue par un automate déterministe complet (Q, q_i, T, F). On considère la substitution α définie sur l'alphabet Q par $\alpha(q) = s_0 \dots s_{k-1}$ avec $s_0 = q$ et $q \xrightarrow{i} s_i$. On considère le point fixe $w = \alpha^\omega(q_i)$. Pour un entier n écrit en base k $n_0 \dots n_p$, la lettre $w_{[n_0 \dots n_j]}$ représente[7] l'état dans lequel se trouve l'automate après avoir lu les $j+1$ premières lettres. On a donc :

$$w_n \in F \Leftrightarrow n \in P$$

Pour la réciproque, on construit l'automate suivant le morphisme selon le schéma montré ci-dessus. ◀

Exemple 5.1. La 2–substitution α définie par $\alpha(a) = ab$ et $\alpha(b) = ba$ a pour point fixe m, le mot de Thue-Morse et l'on vérifie facilement que $\{n, m_n = a\}$ est 2–rationnel.

5.2 Morphismes substitutifs primitifs

Ces morphismes substitutifs donnent lieu à un autre résultat similaire au théorème de Cobham. Ici, les morphismes substitutifs seront à valeurs non pas dans A^k mais dans A^*.

Définition 5.2. *Pour un tel morphisme α on note M_α la matrice carrée définie par $m_{ij} = |\alpha(j)|_i$, c'est-à-dire le nombre d'occurrences de la lettre i dans le mot $\alpha(j)$. Le morphisme α est dit primitif si une puissance de M_α a tous ses coefficients strictement positifs.*

La multiplication de telles matrices est compatible avec la composition des

[7] On note $[n_0 \dots n_j] = \sum_{i=0}^{j} n_i k^{j-i}$

morphismes. Ainsi, d'après le théorème de Perron-Frobenius, un morphisme primitif a une matrice possédant une valeur propre simple et réelle λ, de module strictement plus grand que les autres, possédant un vecteur propre de composantes strictement positives. Le morphisme est alors dit λ*–primitif substitutif.* On a le théorème suivant :

Théorème 5.2. *Soit x un mot infini point fixe de deux morphismes λ– et κ–primitifs substitutifs. Si x n'est pas ultimement périodique alors λ et κ sont multiplicativement dépendants.*

Ce théorème peut d'une certaine manière être vu comme une généralisation du théorème de Cobham. Soit en effet une partie $P \subset \mathbb{N}$ k– et l–rationnelle. Considérons les morphismes substitutifs α_k et α_l définis comme dans la preuve de la proposition 5.1. Les deux points fixes associés, sur les alphabets K et L, peuvent être reliés par un morphisme lettre à lettre φ, pourvu que $\varphi(K_P) \subset \varphi(L_P)$ et $\varphi({}^cK_P) \subset \varphi({}^cL_P)$. On suppose maintenant que les deux morphismes sont primitifs. Alors, ils sont respectivement k– et l– primitifs substitutifs. En effet, la somme des éléments sur une ligne de ${}^tM_K$ est égale à k qui est donc une valeur propre pour le vecteur $(1, \ldots, 1)$. Soit μ une autre valeur propre et $n = |A|$. μ est aussi une valeur propre de ${}^tM_K = \left(m'_{ij}\right)$. Soit $(x_1, \ldots, x_n)$ un vecteur propre associé. Pour tout i on a :

$$\sum_{j=1}^{n} m'_{ij} x_j = \mu x_i$$

En sommant sur i on obtient $k \sum x_i = \mu \sum x_i$. Si μ est la valeur propre dominante alors $\sum x_i \neq 0$ et $k = \mu$. Ceci montre que k est la valeur propre dominante. Le théorème de Cobham donne, dans ce cas, le même résultat que le théorème 5.2.

Références

[1] D. Perrin. Finite Automata. In *Handbook of Theoretical Computer Science*, volume B, chapter 1, pages 1–57. 1990.

[2] F. Durand. Sur les ensembles d'entiers reconnaissables. In *Journal de Théorie des Nombres de Bordeaux*, pages 65–84. 1998.

La référence majeure utilisée ici est la première, qui contient une bibliographie très complète. La deuxième référence concerne la dernière section.