Motifs inévitables

Sommaire

Source :

Cassaigne Julien.
Unavoidable patterns.
Ouvrage collectif de M. Lothaire, "Algebraic Combinatorics on Words", Encyclopedia of Mathematics and its Applications, 90, Cambridge University Press, xiv+504 pp. ISBN: 0-521-81220-8, 2002.

Introduction

On utilise deux alphabets : A l'alphabet ordinaire dont les éléments sont les lettres a,b,c... et sur lequel on construit les mots usuels, et E dont les éléments sont les variables, notées α,β,γ ... Les éléments de E* sont appelés motifs. Le langage associé au motif p est le langage sur A constitué des mots qui peuvent s'écrire $\mu(p)$ où μ est un morphisme non effaçant de E* dans A*. On note ce langage $p(A^+)$.
Par exemple si $p = \alpha\alpha\beta\beta\alpha$ alors $p(A^+) = \{uuvvu \mid u, v \in A^+\}$.
On dit que le mot $w \in A^*$ rencontre le motif p si w a un facteur qui appartient au langage associé à p, sinon on dit que w évite p.
Avec $p = \alpha\alpha\beta\beta\alpha$ alors 1011011000111 rencontre p (avec $\mu : \alpha \to 011, \beta \to 0$) tandis que 0000100010111 évite p.
Un motif p est évitable sur A s'il y a une infinité de mots de A* qui évitent p. Dans le cas contraire, il y a un nombre fini de mots de A* qui évitent p donc il existe un entier n tel que tout mot de longueur supérieure à n rencontre p, on dit que p est inévitable sur A.

Propriété : p est évitable sur A si et seulement si il existe un mot infini sur A* qui évite p.

Preuve : Soit x un mot infini évitant p, alors x n'a aucun facteur dans $p(A^+)$ donc tous les préfixes de x évitent p et leur ensemble est infini. La réciproque utilise le lemme de König : l'arbre des mots de A* qui évitent p (avec pour racine e et où les fils d'un mot w sont du type wa, $a \in A$) est un arbre infini où tout nœud est de degré fini car il a au plus card A fils, il a donc une branche infinie donc il existe un mot infini sur A* qui évite p. ■

Si A est de cardinal k et si p est évitable sur A, alors p est évitable sur tout alphabet de cardinal k et on dit que p est k-évitable.
Plus généralement on dit que p est évitable si il existe un alphabet A tel que p est évitable sur A.
p est dit inévitable si il n'est évitable sur aucun alphabet.

1.Quelques exemples

Si A est unaire, tout motif p est inévitable sur A car tout mot de A* de longueur supérieure ou égale à celle de p rencontre p.

1.1.Puissances d'une variable

$\alpha^0 = \varepsilon$ et $\alpha^1 = \alpha$ sont inévitables car tout mot non vide les rencontre.
$\alpha^2 = \alpha\alpha$ est évitable et plus précisément 3-évitable et 2-non évitable.
α^n =pour $n \geq 3$ est 2-évitable et bien sur 1-inévitable.

Pour prouver ces résultats, on commence par étudier les chevauchements. Un chevauchement est un mot de la forme uvuvu ou u est un mot non vide et v un mot quelconque. Un mot est dit sans chevauchement s'il ne contient pas de facteur qui soit un chevauchement.
La terminologie est justifiee par la remarque suivante. Un mot w contient un chevauchement si et seulement si un de ses facteurs a deux occurrences qui se chevauchent. En effet si w a un chevauchement uvuvu, le mot uvu a deux occurrences qui se chevauchement dans w. Réciproquement, soit w un mot ayant deux occurrences d'un facteur u qui se chevauchent. Soit a la première lettre de u et soit v le reste de u jusqu'au debut de l'occurrence suivante. Le mot avava est un chevauchement de w. Ceci montre en particulier qu'un mot ayant un chevauchement a toujours un chevauchement de la forme avava ou a est une lettre.
On commence par montrer qu'il existe des mots arbitrairement longs sans chevauchement sur l'alphabet A = {0, 1}. Soit le morphisme de Thue-Morse τ de A* dans A* défini de la manière suivante :
$\tau : 0 \to 01$
$\quad 1 \to 10$
Le mot infini $\tau^\omega(0)$ est appelé mot de Thue-Morse. On prouve sans difficulté que la n-ieme lettre du mot de Thue-Morse est 0 si le nombre de 1 dans l'écriture binaire de n est pair et 1 sinon. Le début du mot de Thue-Morse est le suivant : 0110100110010110100101100110100 · · ·

Proposition : Le mot de Thue-Morse est sans chevauchement.

Soit X l'ensemble (01 + 10)*. Par définition même du morphisme τ, tout mot $\tau^n(0)$ appartient à X pour $n \geq 1$.

On commence par établir une propriété élémentaire de l'ensemble X.

Lemme : Si x appartient à X, alors 0x0 et 1x1 n'appartiennent pas à X.

Preuve : On raisonne par récurrence sur la longueur de x. Si x est le mot vide, ni 00 ni 11 n'appartient à X. Soit $x = x_1 \cdots x_n$ un mot de X de longueur $n \geq 1$. Si 0x0 appartient a X, on a $x_1 = x_n = 1$ et le mot $x_2 \cdots x_{n-1}$ appartient a X. Ceci aboutit a une contradiction. Le raisonnement pour 1x1 est similaire.
Preuve de la proposition. Il suffit de montrer que les mots $\tau^n(0)$ sont sans chevauchement. On raisonne par récurrence sur n. Pour n = 1, le mot $\tau(0) = 01$ ne contient pas de chevauchement. Supposons par l'absurde que le mot $w = \tau^{n+1}(0)$ contienne un chevauchement. Comme cela a été vu, on peut supposer que ce chevauchement est de la forme avava ou $a \in \{0, 1\}$ et $v \in \{0, 1\}^*$. Le mot w se factorise donc w = xavavay. On peut supposer que x est de longueur paire car 'autre cas s'obtient en passant au mot miroir puisque x et y ont des longueurs de parités différentes. Si v est de longueur paire, la seconde occurrence de v commence a une position paire. On en déduit que v et ava sont des mots de X* et ceci est une contradiction d'après le lemme précédent. Si v est de longueur impaire, il se factorise v = bv′. Comme le premier et le dernier a de avava sont à des positions de même parité, la première lettre de y est aussi b et y se factorise y = by′. Le mot w se factorise finalement w = xabv′abv′aby′. Comme x est de longueur paire, le mot abv′abv′ab est l'image d'un facteur auaua de $\tau^n(0)$. Ceci est en contradiction avec l'hypothèse de récurrence. ■

On considère le morphisme σ de {0, 1, 2} dans {0, 1} défini de la manière suivante.
$\sigma : 0 \rightarrow 0$
$1 \rightarrow 01$
$2 \rightarrow 011$
Chaque mot $\tau^n(0)$ ne contient pas de facteur 111 et commence par 0. Il appartient donc a l'ensemble (0+01+011)*. On en déduit que pour tout entier n, il existe un mot w_n sur l'alphabet {0, 1, 2} tel que $\sigma(w_n) = \tau^n(0)$. Le mot w_n est en outre unique. Chaque mot w_n est préfixe du mot w_{n+1}. La suite $(w_n)_{n\geq 0}$ converge vers un mot infini que l'on note $\sigma^{-1}(\tau^{\omega}(0))$. Le début de ce mot est le suivant :
210201210120210201202101210201 20 · · ·

Théorème (Thue 1906) : Le mot $\sigma^{-1}(\tau^{\omega}(0))$ est sans carré (c'est à dire qu'il ne rencontre pas le motif αα).

Preuve : Il suffit de prouver que chacun des mots w_n est sans carré. Supposons par l'absurde que w_n contienne un carré uu. On peut supposer que uu a une occurrence qui n'est pas a la fin de w_n. Sinon, on remplace w_n par w_{n+1}. Comme uua est un facteur de w_n, $\sigma(u)\sigma(u)\sigma(a)$ est un facteur de $\tau^n(0)$. Les mots σ(u) et σ(a) commencent par 0. Le mot $\tau^n(0)$ a alors un chevauchement avavacontrairement au résultat de la proposition précédente. ■

Le mot infini $\sigma^{-1}(\tau^{\omega}(0))$ peut être directement obtenu en itérant un morphisme. Soit le morphisme μ défini de la manière suivante.
$\mu : 0 \rightarrow 1$
$1 \rightarrow 20$
$2 \rightarrow 210$
La suite de mots $(\mu^n(2))_{n\geq 0}$ converge vers le mot $\sigma^{-1}(\tau^{\omega}(0))$. En effet, on vérifie l'égalite $\tau \circ \sigma = \sigma \circ \mu$. En utilisant en outre l'égalite $\sigma(2)0 = \tau^2(0)$, on montre par récurrence que $\sigma(\mu^n(2))\tau^n(0) = \tau^{n+2}(0)$.

Le mot $\sigma^{-1}(\tau^{\omega}(0))$ est un mot infini sans carré sur un alphabet de cardinal 3 donc le motif $\alpha^2 = \alpha\alpha$ est 3-évitable. Sur l'alphabet a deux lettres {0, 1}, les seuls mots qui évitent αα sont les mots sans carré soit 010, 101 et leurs facteurs donc le motif $\alpha^2 = \alpha\alpha$ est 2-inévitable.
Le mot de Thue-Morse est un mot infini sans chevauchement sur un alphabet de cardinial 2, or les chevauchements uvuvu correspondent au motif αβαβα si v est non vide et au motif $\alpha^3 = \alpha\alpha\alpha$ si v est vide, donc α^3 est 2-évitable ainsi que αβαβα.

1.2.Autres résultats

Propriété : αβα est inévitable. Plus exactement si le cardinal de A est k, tout mot de longueur au moins 2k+1 rencontre αβα et il existe un mot de longueur 2k qui évite αβα.

Preuve : Soit $w \in A^*$ un mot de longueur au moins 2k+1. Alors une des lettres de A qu'on note a apparaît au moins trois fois dans w. On pose $w = w_0 a w_1 a w_2 a w_3$, h(α) = a et $h(\beta) = w_1 a w_2$. Alors $h : \{\alpha, \beta\}^* \rightarrow A^*$ est un morphisme non effaçant et h(αβα) est un facteur de w donc w rencontre αβα. Si $A = \{a_1 , a_2, ..., a_k\}$, alors

$a_i\, a_1\, a_2\, a_2 \ldots a_k\, a_k$ est un mot de longueur 2k qui évite $\alpha\beta\alpha$. ■

Lemme : Soit p un motif inévitable sur un alphabet A et λ une variable qui n'apparait pas dans p, alors le motif $p\lambda p$ est inévitable sur A .

Preuve : L'ensemble des mots sur A qui évitent p est fini. Il existe un entier m tel que tout mot de longueur m contient le motif p. Soit $n = |A|^m$, le nombre de mots sur A de longueur m. Un mot w de longueur nm + m + n se factorise $w = u_0a_1u_1a_2 \cdots a_nu_n$ ou $a_1, \ldots, a_n$ sont des lettres et $u_0, \ldots, u_n$ sont des mots de longueur m. Par définition de n, il existe deux mots u_i et u_j pour i < j qui sont égaux. Le mot w se factorise donc $v_0u_iv_1u_iv_2$. Comme u_i est de longueur m, il se factorise $u_i = v_3\mu(p)v_4$ ou μ est un morphisme de $(E \setminus \{\lambda\})^*$ dans A^*. En prolongeant μ à E^* par $\mu(\lambda) = v_4v_1v_3$, le mot w contient le facteur $\mu(p\lambda p)$. ■

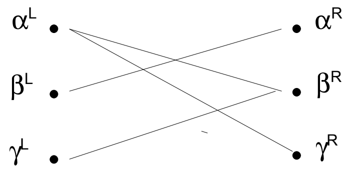
Mots de Zimin : Soit X l'alphabet infini $\{x_0, x_1, x_2, \ldots\}$. On définit la suite $(z_n)_{n\geq 0}$ des mots de Zimin par $z_0 = \varepsilon$ et $z_{n+1} = z_nx_nz_n$ pour tout $n \geq 0$. On montre par récurrence sur n que chacun des mots z_n est un motif inévitable : on a déjà vu que $z_0 = \varepsilon$, $z_1 = \alpha$ et $z_2 = \alpha\beta\alpha$ sont inévitables. Enfin si z_n est inévitable, z_{n+1} est inévitable d'après le lemme précédent. La suite des mots de Zimin converge vers un mot infini z. Pour tout $n \geq 0$, la n-ieme lettre de z est x_k si 2k est la plus grande puissance de 2 qui divise n. Le début du mot z est le suivant :
$z = x_0x_1x_0x_2x_0x_1x_0x_3x_0x_1x_0x_2x_0x_1x_0x_4x_0x_1x_0x_2x_0x_1x_0x_3 \cdots$

On a donné à chaque fois une preuve s'adaptant au cas particulier étudié. Il existe une façon plus systématique de montrer qu'un motif donné est évitable ou non, par l'algorithme de Zimin.
Par contre, aucun algorithme n'est connu pour déterminer si un mot est k-inévitable pour un entier k donné.

2. L'algorithme de Zimin

2.1. Irréductibilité

Le graphe d'adjacence du motif p est le graphe AG(p) comprenant comme sommets deux copies verticales de E, une à gauche (L) et une à droite (R) notées E^L et E^R. Il y a une arrête entre κ^L et λ^R si et seulement si $\kappa\lambda$ est un facteur de p. Par exemple voici $AG(\alpha\beta\alpha\gamma\beta\alpha)$, qui a six sommets et quatre arrêtes.



Si F est une partie de E on note δ_F le morphisme de E^* dans $(E\setminus F)^*$ qui associe ε à une variable de F et laisse les variables de $E\setminus F$ inchangées. Ce morphisme efface les variables qui sont dans F.
Une partie F de E est dite libre pour p si il n'existe pas de chemin dans AG(p) reliant κ^L et λ^R avec κ et λ dans F. Dans l'exemple il y a deux parties libres, $\{\alpha\}$ et $\{\beta\}$. Par contre $\{\gamma\}$ n'est pas libre car il y a un chemin de γ^L à γ^R . Pour un motif p, on dit que p se réduit à q en une étape si il existe F une partie libre pour p telle que $q = \delta_F(p)$. On dit que p est réductible s'il se réduit à ε en un nombre fini d'étapes. Sinon, p est irréductible.
Dans notre exemple, $p = \alpha\beta\alpha\gamma\beta\alpha$ est réductible : il se réduit à $\beta\gamma\beta$ par suppression de la partie libre $\{\alpha\}$, et $\beta\gamma\beta$ se réduit à γ par supression de la partie $\{\beta\}$ libre pour $\beta\gamma\beta$, et ce dernier se réduit à ε .Par contre si on avait commencé par réduire $\alpha\beta\alpha\gamma\beta\alpha$ par supression de la partie libre $\{\beta\}$, on aurait obtenu le motif irréductible $\alpha\alpha\gamma\alpha$. Ainsi pour prouver qu'un motif est irréductible, il faut essayer récursivement toutes les possibilités de réduction, donc l'algorithme n'est pas très efficace en pratique. Cet algorithme est appelé l'algorithme de Zimin.
Notations : Etant donnés un motif p et son graphe d'adjacence AG(p), soit X un ensemble de sommets de AG(p), on note C(X,p) l'ensemble des sommets de AG(p) qui sont reliés à un élément de X par un chemin dans le graphe. On note $C_L(X,p)$ l'ensemble des variables $\lambda \in E$ telles que $\lambda^L \in C(X,p)$ et on définit de même $C_R(X,p)$. F est une partie libre pour p si et seulement si $F \subset C_L(F^L,p)\setminus C_R(F^L,p)$

2.2 Lien entre motif irréductible et évitable

Theorème : Un motif est décitable si et seulement si il est irréductible.

On peut donc décider si un motif est inévitable ou non grâce à l'algorithme de Zimin.
La démonstration est longue et fait intervenir plusieurs lemmes.

2.2.1. Si un motif est réductible alors il est inévitable.

Lemme : Si p se réduit à q et si q est inévitable, alors p est inévitable.

Si on a le lemme, comme ε est inévitable et qu'un motif réductible se réduit à ε , en prenant q = ε on a le résultat voulu. Il suffit donc de démonter le lemme.

Preuve : On va montrer que si p se réduit à q en une étape et si q est inévitable, alors p est inévitable. On aura alors le lemme en raisonnant par récurrence sur le nombre d'étapes de la réduction. Soit un motif p qui se réduit à un motif inévitable q en une étape, donc par supression d'une partie F libre pour p. Montrons que p est inévitable sur tout alphabet A par récurrence sur le cardinal de A.
Si Card A = 1, la propriété est vraie car tout motif est inévitable sur A. Si A = B U {a} et si p est inévitable sur B (hypothèse de récurrence), l'ensemble des mots sur B évitant p est $L = p(B^+)\backslash B^*p(B^+)B^*$ qui est fini par hypothèse. Soit $M = aA^* \backslash A^*p(A^+)A^*$ l'ensemble des mots sur A évitant p et commencant par a. Chaque mot de M qui n'est pas une puissance de a peut s'écrire comme produit non vide de mots de $N = \{a^iwa^j \mid w \in L, 0<i< |p|, 0 \leq j < |p|\}$, qui est fini car L l'est. On peut voir N comme un nouvel alphabet, soit i le morphisme de N^* dans A^* qui à un élément de N (vu comme une lettre) associe lui-même vu comme un mot sur A. Alors $M \subset i(N^+) \cup a^+$. Soit λ une variable qui n'apparait pas dans p, alors λ n'apparait pas dans q qui est inévitable et d'après une propriété vue en I, $q\lambda q$ est inévitable, donc $q\lambda$ est inévitable. Soit un mot suffisament long $w \in N^*$, il existe donc un morphisme non effaçant µ de E^* dans N^* tel que $\mu(q\lambda)$ est un facteur de w. On remarque que pour toute variable κ, $\mu(\kappa) \in N^+$ car µ est non effaçant, donc $i(\mu(\kappa)) \in aA^+$. Soit g un morphisme de E^* dans A^* défini par :
_$g(\kappa) = i(\mu(\kappa))$ si $\kappa \in E \setminus (C_L(F^L,p) \cup C_R(F^L,p))$
_$g(\kappa) = a^{-1}i(\mu(\kappa))$ si $\kappa \in C_R(F^L,p) \setminus C_L(F^L,p)$
_$g(\kappa) = i(\mu(\kappa))a$ si $\kappa \in C_L(F^L,p) \setminus (C_R(F^L,p) \cup F)$
_$g(\kappa) = a^{-1}i(\mu(\kappa))a$ si $\kappa\ C_L(F^L,p) \cap C_R(F^L,p)$
_$g(\kappa) = a$ si $\kappa \in F$
Notons que g est bien défini car les cinq cas s'excluent mutuellement : F est libre donc $F \subset C_L(F^L,p) \setminus C_R(F^L,p)$, de plus g est non effaçant. De plus si on montre que g(p) est un facteur de $i(\mu(q\lambda))$, alors i(w) rencontre p car $\mu(q\lambda)$ est un facteur de w, donc i(w) n'appartient pas à M. Alors M est fini et p est inévitable sur A .
Il reste donc à montrer que g(p) est un facteur de $i(\mu(q\lambda))$. On va montrer par récurrence sur k, $1\leq k < |p|$, que si p_k est le préfixe de longueur k de p et si p_k se réduit à q_k par supression de F, où q_k est un suffixe de q, alors $rg(p_k) = i(\mu(q_k))s_k$ où r est a ou ε suivant si la premiere lettre de p est dans $C_R(F^L,p)$ ou pas et s_k est a ou ε suivant si la dernière lettre de p_k est dans $C_L(F^L,p)$ ou pas. Pour k = 1 on a le résultat par définition de g. Si et $p_{k+1} = p_k\eta$. La dernière lettre ξ de p_k est telle qu'il y ait un chemin de ξ^L à η^R dans AG(p). Il faut montrer que $rg(p_{k+1}) = i(\mu(q_{k+1}))s_{k+1}$, avec $s_{k+1} = a$ si $\eta \in C_L(F^L,p)$, $s_{k+1} = \varepsilon$ sinon. Or $rg(p_{k+1}) = rg(p_k)g(\eta) = i(\mu(q_k))s_kg(\eta)$ qui se réduit à $s_kg(\eta) = i(\mu(\eta))s_{k+1}$ si $\eta \notin F$, ou a $s_kg(\eta) = s_{k+1}$ si $\eta \in F$. Ceci est contradictoire avec la définition de g car $s_k = a$ si et seulement si $\eta \in C_R(F^L,p)$, car c'est équivalent à $\xi \in C_L(F^L,p)$. ■

2.2.2. Si un motif est inévitable, alors il est réductible.

Lemme 1 : Soient p et q des motifs tels que f(q) est un facteur de p pour f morphisme non effaçant de E*, et F une partie libre pour p. Soit F' l'ensemble des variables $\xi \in E$ telles que $f(\xi) \in F^+$. Alors F' est une partie libre pour q. De plus si p se réduit à p' par supression de F et si q se réduit à q' par supression de F', alors f '(q') est un facteur de p' où $f' = \delta_F \circ f|_{E \setminus F'}$ est le morphisme non effaçant de $(E\setminus F')^*$ dans $(E\setminus F)^*$ qui associe à une variable ξ le motif obtenu en suprimant les éléments de F dans $f(\xi)$.

Preuve : Soit une variable $\xi \in E$, on note $f_1(\xi)$ la première lettre de $f(\xi)$ et $f_2(\xi)$ la dernière lettre de $f(\xi)$. Si $\xi\eta$ est un facteur de q, alors $f_2(\xi)f_1(\eta)$ est un facteur de p. Si on remplace ξ^L par $f_2(\xi)^L$ et ξ^R par $f_1(\xi)^R$, on voit que AG(q) est un sous-graphe de AG(p). Soient ξ et η dans F', s'il y avait un chemin de ξ^L à η^R dans AG(q), il y aurait un chemin de $f_2(\xi)^L$ à $f_1(\eta)^R$ dans AG(p), ce qui est exclu car $f_2(\xi)$ et $f_1(\eta)$ sont dans F qui est libre pour p. Donc F' est libre pour q, et le reste du lemme est évident. ■
On remarque que si $\xi \in F'$ est un facteur de q, alors $f(\xi)$ est de longueur 1 car deux variables de F (qui est libre pour p) ne peuvent pas être consécutives dans p.

Lemme 2 : Si p, p' et q sont des motifs tels que p se réduit à p' et f(q) est un facteur de p avec f morphisme non effaçant, alors il existe un motif q' et un morphisme non effaçant f ' tels que q se réduit à q' et f '(q') est un facteur de p', vérifant de plus la condition f (alph q \ alph q') $\subset$ (alph p \ alph p')*, où on note alph p l'ensemble des lettres de p, c'est à dire que si ξ est une variable qu'on supprime de q pour former q', alors f(ξ) ne contient que des variables qu'on supprime de p pour former p').

Preuve : Il s'agit d' itérer le lemme 1. On raisonne par induction sur le nombre d'étapes dans la réduction de p à p'. Si p = p', alors le résultat est clair avec q' = q et f ' = f. Si le résultat est vrai pour un nombre n d'étapes et si p se réduit à p' en n+1 étapes, alors p se réduit à p'' en n étapes et p'' se réduit à p' en une étape par supression d'une partie F libre pour p''. Par hypotèse de récurrence, on peut trouver q'' et f '' tels que q se réduit à q'' et f ''(q'') est un facteur de p'', vérifant de plus la condition f (alph q \ alph q'') $\subset$ (alph p\ alph p'')*. On applique alors le lemme 1 à p'' et q'' pour obtenir une partie F' libre pour q'', un motif q' et un morphisme non effaçant f ' tels que q'' se réduit à q' par supression de F', f '(q') est un facteur de p'. On a donc bien l'existence de q' et f ', et la condition est vérifiée car alph q \ alph q' = (alph q \ alph q'') U F or par hypothèse de récurrence f(alph q \ alph q'') $\subset$ (alph p \ alph p'')* $\subset$ (alph p \ alph p')* et f''(F') $\subset$ F^+ donc δ_F(f''(F')) $\subset$ $\delta_F(F^+) \subset \{\varepsilon\}$, d'où f(F') $\subset$ (alph p \ alph p')*. ∎

Lemme 3 : Soit q = δ_v(p) : le motif q est obtenu à partir de p en suprimant les éléments de V, qui n'est pas forcément une partie libre pour p. S'il existe un motif r et un morphisme non effaçant f tels que r se réduit à q et f(p) est un facteur de r, et si $\xi \in V$ si et seulement si f(ξ) $\in$ (alph r \ alph q)*, alors p se réduit à q.

Preuve : On applique le lemme 2 avec r, q et p prenant respectivement les rôles de p, p' et q. Il existe un motif q' et un morphisme non effaçant f ' tels que p se réduit à q' et f'(q') est un facteur de q et f (alph p \ alph q') $\subset$ (alph r \ alph q)*. Or f(ξ) $\in$ (alph r \ alph q)* si et seulement si $\xi \in V$, donc (alph p \ alph q') $\subset$ V d'où δ_V(q') = δ_V(p) = q. Mais f'(q') est un facteur de q avec f' non effaçant donc |q'|≤|q| et δ_V(q') = q donc q' = q et p se réduit à q. ∎

Soit k un entier naturel, on considère l'alphabet à 4k-lettres $A_k = \{a_0, a_1, ..., a_{2k-1}, b_0, b_1, ..., b_{2k-1}\}$. On définit le morphisme φ_k sur A_k par $\varphi_k(a_i) = a_0 b_i a_1 b_{i+1} ... a_{k-1} b_{i+k-1}$ et $\varphi_k(b_i) = a_k b_i a_{k+1} b_{i+1} ... a_{2k-1} b_{i+k-1}$ où on considère les indices modulo 2k. Le morphisme est uniforme : l'image de toute lettre à pour longueur 2k. A partir de φ_k on construit le mot infini $w^{(k)} = \varphi_k^{\omega}(a_0)$ qui est le point fixe de φ_k. On va montrer que tout motif irréductible est évité par $w^{(k)}$ quel que soit k.

Lemme 4 : Soit v un facteur de longueur au moins 2 de $w^{(k)}$. Alors il existe un entier i, $0 \le i \le 4k-1$et une lettre $x \in A_k$ telle que à chaque fois que v apparaît dans $w^{(k)}$ à une position n, alors n = i (mod 4k) et la lettre à la position n' = $\left[\frac{n}{2k} \right]$ dans $w^{(k)}$ est $w^{(k)}_{n'} = x$.

Preuve : On suppose que |v| = 2, le cas général s'en déduit facilement. On remarque que dans une image par φ_k, donc dans $w^{(k)}$, les a et les b s'alternent. Or $w^{(k)} = \varphi_k(w^{(k)})$ donc $w^{(k)} = \varphi_k(a_p)\, \varphi_k(b_q)\, \varphi_k(a_r)\, \varphi_k(b_s)\, \varphi_k(a_t)\, \varphi_k(b_u)$... et vu la forme des $\varphi_k(a_i)$ et $\varphi_k(b_i)$, on a $w^{(k)}_{2i} = a_i$, où l'indice i est considéré modulo 2k. Donc les lettres a_i permettent de connaître la position dans $w^{(k)}$ modulo 4k. Chaque facteur de longueur 2 de $w^{(k)}$ contient au moins un a_i donc la position de v est unique modulo 4k. Si il y a une occurrence de v à la position n dans $w^{(k)}$ alors la premère lettre de v fait partie de l'image de x = $w^{(k)}_{n'}$. Donc soit v est un facteur de φ_k(x), soit v est formé de la dernière lettre de φ_k(x) et de la première lettre de l'image du successeur de x, c'est à dire a_0 ou a_k. On est alors dans l'un des 4 cas suivants :

- Si v = $a_i b_j$ avec $0 \le i \le k-1$, alors x = a_{j-i}
- Si v = $b_j a_{i+1}$ avec $0 \le i \le k-1$, alors x = a_{j-i}
- Si v = $a_i b_j$ avec $k \le i \le 2k-1$, alors x = b_{j+k-1}
- Si v = $b_j a_{i+1}$ avec $k \le i \le 2k-1$, alors x = b_{j+k-1}

Donc x, la lettre à la position n', est uniquement déterminée par v. ∎

Lemme 5 : Soit un motif p, k un entier tel que 2k > Card(alph p) et v un facteur de $w^{(k)}$ tel que φ_k(v) rencontre p. Alors il existe un motif q tel que p se réduise à q et v rencontre q.

Preuve : Soit h un morphisme non effaçant tel que h(p) soit un facteur de φ_k(v). D'après le lemme 4, pour chaque lettre $x \in A_k$ il existe 2k mots de longueur 2, facteurs de φ_k(x) ou formé de la dernière lettre de φ_k(x) et de la première lettre de l'image du successeur de x, qui permettent de déterminer x. Ces 2k mots terminent tous par des lettres différentes, or 2k > Card(alph p) donc au moins l'un de ces mots termine par une lettre qui n'est pas la première lettre de h(ξ) pour un ξ dans alph p. On choisit un tel mot et on le note d_x, mot caractéristique pour x. Par construction, à chaque fois que d_x apparaît dans h(p), il est facteur d'un h(ξ). Soit V l'ensemble des variables ξ telles que h(ξ) ne contienne pas de mot caractéristique, et q = δ_v(p). On définit h' morphisme non effaçant de (E\V)* dans A_k* par h'(ξ) = $x_1 x_2 ... x_m$ où $d_{x1} d_{x2} ... d_{xm}$ sont les mots

caractéristiques qui apparaîssent dans $h(\xi)$, dans leur ordre d'apparition. Chaque lettre x de h'(q) est une lettre de v : chaque mot caractéristique qui apparaît dans un $h(\xi)$ apparaît dans h(p) qui est un facteur de $\varphi_k(v)$, or un mot caractéristique d_x est un facteur (ou assimilé) de $\varphi_k(x)$, donc x est une lettre de v. De plus des mots caractéristiques consécutifs correspondent à des lettres consécutives de v, donc h'(q) est un facteur de v, c'est à dire v rencontre q.
Il reste à montrer que p se réduit à q. Comme $q = \delta_v(p)$ on va utiliser le lemme 3. On commence par construire un morphisme non effaçant f tel que f(p) se réduit à q. On définit f de E* dans (E U A)* ainsi :

- Si $\xi \in V$ alors $f(\xi) = h(\xi)$
- Si $h(\xi)$ ne contient ni a_0 ni a_k, mais contient un mot caractéristique d_x (il ne peut contenir qu'un seul mot caractéristique dans ce cas) alors $\varphi_k(x) = a_i v_1 h(\xi) v_2$, avec $i \in \{0,k\}$ et $v_1, v_2 \in A_k^*$ et on pose $f(\xi) = h(\xi) v_2 \xi v_1 h(\xi)$
- Si $\xi \in V$ et $h(\xi) = v_1 \varphi_k(w) a_j v_2$, avec $j \in \{0,k\}$ et $v_1, v_2, w \in A_k^*$, avec w de longueur maximale, alors soit a_l la première lettre de $\varphi_k(w)a_j$ (donc $i \in \{0,k\}$), x_1 la première lettre de $h'(\xi)$ et x_2 la dernière lettre de $h'(\xi)$. On pose $f(\xi) = v_1 v'_1 \xi v'_2 v_2$ où $v'_1 = \varepsilon$ si le premier mot caractéristique d_{x1} de $h(\xi)$ est facteur de $v_1 a_l$, $v'_1 = \varphi_k(x_1)$ sinon, et de même $v'_2 = \varepsilon$ si le dernier mot caractéristique de $h(\xi)$ est facteur de $a_j v_2$, $v'_2 = (a_{j+k})^{-1} \varphi_k(x_2) a_j$ sinon.

En supprimant les éléments de A_k de f(p), on obtient q car $f(\xi) \in A_k^*$ si et seulement si $\xi \in V$ et sinon $f(\xi) \in A_k^* \xi A_k^*$ donc $\delta_A(f(p)) = \delta_v(p) = q$. De plus deux variables ξ_1 et ξ_2 consécutives dans q sont séparées dans f(p) par le mot $v_1 a_i v_2 \in A_k^*$, où $a_{i+k} v_1$ est l'image par φ_k de la dernière lettre de $h'(\xi_1)$ et $a_i v_2$ est l'image par φ_k de la première lettre de $h'(\xi_2)$. Ceci permet de réduire f(p) ainsi : d'abord $F_0 = \{b_0, b_1,..., b_{2k-1}\}$ est une partie libre pour f(p) car un élément de F_0 peut uniquement être suivi d'un élément de E U $\{a_0, a_1 , ..., a_{2k-1}\}$ qui ne peut être précédé que par un élément de F_0. Donc par supression de F_0
, f(p) se réduit à p_0 où p_0 ne contient que des variables de E et des lettres a_i, qui apparaîssant par séquences $a_1 a_2 ...a_{k-1} a_k a_{k+1}...a_{2k-1}$ ou $a_{k+1}a_{k+2}...a_{2k-1}a_0 a_1 ...a_{k-1}$ entre deux variables de E. Alors $F_1 = \{a_1, a_{k+1}\}$ est une partie libre pour et on peut réduire p_0 à p_1 . On continue en suprimant à chaque fois $F_i = \{a_i, a_{k+i}\}$ pour i de 2 à k-1, et on obtient un motif $p_{k-1} \in (E \cup \{a_0, a_k\})^*$ dans lequel les éléments de E et $F_k = \{a_0, a_k\}$ s'alternent, donc F_k est libre pour p_{k-1} et p_{k-1} se réduit à q. ■

Lemme 6 : Si le mot infini $w^{(k)}$ rencontre un motif p contenant moins de 2k variables distinctes, alors p est réductible.

Preuve : Il existe un entier m tel que $\varphi_k^m(a_0)$ rencontre p. On applique le lemme 5 et on obtient un motif p_1 tel que p se réduise à p_1 et $\varphi_k^{m-1}(a_0)$ rencontre p_1 . On réitère le processus car on a toujours 2k > Card(alph p_i). On obtient m motifs p_i pour $1 \le i \le m$ tels que p se réduise à p_i et $\varphi_k^{m-i}(a_0)$ rencontre p_i. Mais alors a_0 rencontre p_m ce qui signifie que p_m est soit une variable, soit ε et dans tous les cas p_m est réductible donc p aussi. ■

preuve du théorème : On a déjà montré que si p est réductible, alors p est inévitable. Inversement supposons que p est inévitable, alors tout mot infini rencontre p et en particulier pour $k = \left\lceil \frac{(\mathrm{Card\,alph}\,p+1)}{2} \right\rceil$ alors $w^{(k)}$ rencontre p et d'après le lemme 6 p est réductible. ■

Corollaire 1 : Le mot infini $w^{(k)}$ évite tous les motifs évitables qui ont au plus 2k-1 variables.

Preuve : Si $w^{(k)}$ rencontre un motif p ayant au plus 2k-1 variables, alors par le lemme 6 p est réductible donc p est inévitable. ■

Corollaire 2 : Si chaque variable d'un motif p apparaît au moins 2 fois dans p, alors p est évitable.

Preuve : Si p est inévitable, alors il peut se réduire à une variable α donc α est une variable qui n'est qu'une seule fois dans p. ■

Corollaire : Soit p un motif de n variables distinctes. Si $|p| \ge 2^n$ alors p est évitable.

Preuve : On montre par récurrence sur n que si p est inévitable, alors $|p| < 2^n$. Pour n = 1, le résultat est vrai car les carrés sont évitables. Si le résultat est vrai pour n alors soit p un motif inévitable de n+1 variables. Vu le corollaire 2, il existe une variable α qui n'est qu'une fois dans p. Donc p s'écrit $p_1 \alpha p_2$ où p_1 et p_2 sont des motifs avec au plus n variables distinctes qui sont tous deux inévitables car p l'est. Par hypothèse de récurrence, p_1 et p_2 sont au plus de longueur 2^n-1 donc p est au plus de longueur 2^{n+1}-1. ■