Motifs Inévitables

Martin Clochard

16 décembre 2010

Table des matières

1 Définitions et exemples

L'inévitablité d'un ensemble de mots se définit de manière intuitive par le fait que dans tout mot suffisamment grand, on trouve un de ces mots en facteur. Un motif inévitable est un cas particulier où cet ensemble est défini par un motif, par exemple le motif $\alpha\beta\alpha$ qui donne l'ensemble des mots de cette forme, ou l'ensemble des carrés. Formalisons maintenant cette notion.

1.1 Motifs et notions d'évitablité

Définition 1.1.1. Soit A, E deux alphabets finis. E est appelé *l'alphabet des variables.* Un morphisme μ de E^* dans A^* est dit *non-effaçant* si $\forall a \in A, \mu(a) \neq \epsilon$. Un *motif* sur E est simplement un mot sur E. Le langage associé à un motif p dans A est $p(A^+) = \{\mu(p) | \mu$ morphisme non-effaçant de E^* dans $A^*\}$. Il s'agit tout simplement du langage où on a remplacé les lettres de p, les variables, par tout les mots non vides possibles de A^*.

Conventions :
Dans la suite, E représentera l'alphabet des variables et A un alphabet fini quelconque. Les mots écrits avec des lettres grecques seront considéré comme des motifs. La notation $alph(p)$ se réfère à l'alphabet de définition de p, où p est un motif.

Exemple 1.1.2. L'exemple le plus courant de motif est le carré $\alpha\alpha$, dont le langage associé est l'ensemble des mots carrés.

Définition 1.1.3. On dit qu'un mot ω *évite* le motif p si $\exists u \in p(A^+)$:
u est facteur de ω. Un motif $p \in E^*$ est dit *inévitable* sur A si $\exists n \in \mathbb{N}$:
$\forall \omega \in A^*, |\omega| \geq n \Rightarrow \omega$ n'évite pas p, autrement dit si tout mot suffisamment long de A contient un facteur dans le langage associé à a p. Il est dit évitable sinon.
La notion d'évitabilité dépend donc de la taille de l'alphabet. On dit alors que
- Un motif p est inévitable s'il est inévitable quelque soit la taille de l'alphabet, évitable s'il existe un alphabet qui évite p.
- Un motif est k-inévitable s'il est inévitable pour un alphabet de taille k (et donc pour tout les alphabets de taille inférieure), k-évitable sinon.

Exemple 1.1.4. Prenons le motif $\alpha\beta\alpha\alpha$: Le mot *baabaaababbaab* n'évite pas ce motif, puisque il contient $(ab)(aa)(ab)(ab)$ comme facteur.
Le mot *bbbababa*, par contre, l'évite.

Remarque 1.1.5. Tout les motifs sont inévitables sur un alphabet de taille 1.

Proposition 1.1.6. *Un motif p est inévitable sur A si et seulement si il existe un mot infini de A qui évite p.*

Démonstration. Le sens réciproque est évident. Pour le sens direct, l'ensemble des mots de A^* qui évitent p forme un arbre pour la relation $u \to v \Leftrightarrow \exists a \in A : ua = v$. Or, cet arbre est de degré borné par $|A|$. D'après le lemme de König, s'il est infini il a une branche infinie, qui nous donne le mot infini cherché.

□

Définition 1.1.7. Soit p, q deux motifs. On dit que $p|q$ si q n'évite pas p. Cette relation définit évidemment un quasi-ordre sur les motifs, et deux motifs sont alors équivalents ($p|q$ et $q|p$) si et seulement si il existe une permutation des variables qui envoie l'un sur l'autre.

Exemple 1.1.8. Le motif $\alpha\alpha$ divise le motif $\gamma\beta\alpha\beta\alpha$.

Remarque 1.1.9. Si $p|q$ et q est inévitable sur A, alors p aussi.

1.2 L'exemple des puissances

Les premiers motifs qui viennent naturellement à l'esprit sont les puissances, c'est-a-dire les motifs de la forme α^n. Les motifs $\alpha^0 = \epsilon$ et α sont bien sûr inévitables. Par contre, dès $n = 2$ la situation change : le motif $\alpha\alpha$ est 3-évitable (et trivialement 2-inévitable),et le motif $\alpha^n, n \geq 3$ est 2-évitable. On va commencer par établir un lemme utile pour la fabrication de mots infinis évitant un motif.

Lemme 1.2.1. *Soit ϕ un morphisme non-effaçant de A^* dans lui-même. Si a est prefixe propre de $\phi(a)$, alors ϕ admet un unique point fixe commençant par a, qui est la "limite" de $(\phi^n(a))_{n\in\mathbb{N}}$ lorsque $n \to \infty$, c'est-à-dire que les $(\phi^n(a))_{n\in\mathbb{N}}$ sont préfixes propres les uns des autres.*

Démonstration. La preuve se fait par récurrence. l'initialisation est donnée par l'hypothèse.
Hérédité : soit $n \in \mathbb{N}$ tel que la propriété soit vraie. Alors $\phi^{n+1}(a) = \phi(\phi^n(a))$ est préfixe propre de $\phi(\phi^{n+1}(a))$ car $\phi^n(a)$ est préfixe propre de $\phi^n(a)$ par hypothèse de récurrence, ce qui est la propriété qu'il fallait démontrer. On obtient bien un point fixe infini car la longueur des mots de la suite est strictement croissante.
Il reste l'unicité d'un tel fixe ω : Comme il commence par a,tout les $\phi^n(a)$ sont préfixe de ω (les préfixes de ω sont évidemment stables par ϕ). Or le point fixe limite des $(\phi^n(a))_{n\in\mathbb{N}}$ est donc préfixe de ω, soit est ω.

□

Définition 1.2.2. Le mot de *Thue-Morse* est l'unique point fixe commençant par a du morphisme $\Theta : a \to ab$, $b \to ba$, qui existe et est bien défini par le lemme 1.2.1

Proposition 1.2.3. *Le mot de Thue-Morse t évite les motifs $\alpha\beta\alpha\beta\alpha$ et α^3, c'est-à-dire qu'il est sans chevauchement.*

Démonstration. Par l'absurde, supposons qu'il y ait une telle occurrence d'un de ces deux motifs. Considéront une occurrence de taille minimale dans t, soit deux mots $(u, v) \in A^+ \times A^*$ tels que *uvuvu* soit facteur de taille minimale de cette forme dans t. En appliquant Θ une fois à t, on obtient que t est composé de facteurs *ab* et *ba*. Comme *aaa* et *bbb* ne peuvent donc être facteurs de t, cette longueur est supérieure à 5. Comme en particulier tout sous-facteur de t de longueur au moins 5 contient *aa* ou

bb comme facteur d'après la même remarque, $uvuvu$,puis uvu admet aa ou bb comme facteur. Or, toujours d'aprè la remarque ci-dessus, ce facteur ne peut apparaître qu'a des positions impaires dans t. Comme il apparaît dans uvu en position n, il apparaît dans $uvuvu$ en positions n et $n + |uv|$. Donc $|uv|$ est pair. Si $uvuvu$ apparaît à une position paire dans t, on considère les mots u' et v' formés des lettres de u et v a une position paire dans t (Si cette position eétait impaire, on prendrait celles en position impaires et on inverserait a et b). Alors on remarque que $u'v'u'v'u'$ est facteur de t, car si ω est le facteur de t dont l'image par Θ donne le facteur $uvuvux$ (ou x est ϵ ou une lettre pour rectifier la longueur, celle qui suit $uvuvu$ dans t), $\omega = u'v'u'v'u'$ par injectivité de Θ (si les positions étaient impaires, il faudrait rajouter une lettre avant $uvuvu$). C'est absurde par minimalité.

□

Ce résultat démontre notamment que le motif cube, et donc bien sûr toutes les puissances supérieures, sont 2-évitable. Il donne également un résultat pour le motif carré.

Proposition 1.2.4. *Soit μ le morphisme de $\{a, b, c\}^*$ dans $\{a, b, c\}^*$ défini par $\mu : a \to abc$, $b \to ac$, $c \to b$. Alors l'unique point fixe infini commençant par a de μ, noté u, évite le motif carré.*

Démonstration. Soit ν le morphisme de a, b, c^* dans a, b^* défini par $\nu : a \to abb$, $b \to ab$, $c \to a$. Les morphismes $\nu \circ \mu$ et $\Theta \circ \nu$ coïcident sur les lettres de a, b, c, donc sont égaux. Donc $\Theta(\nu(u)) = \nu(\mu(u)) = \nu(u)$. Comme $\nu(u)$ commence par a, le lemme 1.2.1 donne que $\nu(u) = t$. Supposons maintenant que u admette un facteur carré vv, $v \neq \epsilon$. Alors $\nu(vv)$ est facteur de t. Or, par définition de ν,la lettre suivant $\nu(v)$ dans t ne peut être qu'un a : $\nu(v)\nu(v)a$ est facteur de t. De même, $\nu(v)$ commence par un a : $\nu(v) = a\omega \Rightarrow a\omega a\omega a$ est facteur de t, ce qui est absurde par proposition 1.2.3.

□

1.3 Un exemple de motifs inévitables : les motifs de Zimin

On a pour l'instant vu trè peu de motifs inévitables. On va dans cette partie en construire une famille infinie. On va pour cela s'appuyer sur la proposition suivante.

Proposition 1.3.1. *Soit p un motif inévitable et ξ une variable qui n'apparaît pas dans p. Alors $p\xi p$ est inévitable.*

Démonstration. Par définition 1.1.3, $\exists N \in \mathbb{N} : \forall \omega \in A^*, |\omega| \geq N, \omega$ n'évite pas p. Soit $l = |alph(p)|$. Posons $M = |\{\omega \in p(A^+) | |\omega| \leq N\}| + 1$. Soit $\omega \in A^*$ tel que $|\omega| \geq (N + 1)M$. Alors $\omega = u_1 a_1 u_2 a_2 ... u_M a_M v$, avec $(a_i)_{i \in [|1;M|]}$ des lettres et $(u_i)_{i \in [|1;M|]}$ des mots de longueur N. Or par définition de N, chacun des u_i a un facteur dans $p(A^+)$. Par le principe des tiroirs, deux mots u_i et u_j, $i < j$, on un même facteur commun dans $p(A^+)$, soit $\mu(p)$ avec μ un morphisme non effaçant de $alph(p)$ dans A^*. On prolonge μ à $alph(p) \cup \xi$ par $xi \to u_{i+1} ... u_{j-1}$: $\mu(p\xi p)$ est alors facteur de ω. Donc $p\xi p$ est inévitable.

□

Définition 1.3.2. On note $(\alpha_n)_{n \in \mathbb{N}}$ une suite de variables distinctes. Les *motifs de Zimin* sont définis par :
$Z_0 = \epsilon, \forall n \in \mathbb{N}, Z_{n+1} = Z_n \alpha_n Z_n$
Le fait que ces motifs soient inévitables est un corollaire immédiat de la proposition 1.3.1.

Exemple 1.3.3. Les premiers motifs de Zimin sont :

$$Z_0 = \epsilon,\ Z_1 = \alpha,\ Z_2 = \alpha\beta\alpha,\ Z_3 = \alpha\beta\alpha\gamma\alpha\beta\alpha,\ Z_4 = \alpha\beta\alpha\gamma\alpha\beta\alpha\delta\alpha\beta\alpha\gamma\alpha\beta\alpha$$
$$Z_5 = \alpha\beta\alpha\gamma\alpha\beta\alpha\delta\alpha\beta\alpha\gamma\alpha\beta\alpha\zeta\alpha\beta\alpha\gamma\alpha\beta\alpha\delta\alpha\beta\alpha\gamma\alpha\beta\alpha$$

2 L'algorithme de Zimin : décidabilité de l'évitabilité

La preuve de la décidabilité de l'évitabilité d'un motif se fait en en exhibant l'algorithme. Nous allons avoir besoin de la notion de réduction. En effet, la réductibilité au mot vide sera équivalente à l'évitabilité, et surtout déterminable algorithmiquement.

2.1 Réduction de motifs

Définition 2.1.1. Soit F un sous-ensemble de E On appelle δ_F le morphisme de E^* dans $(E \backslash F)^*$ défini par la suppression des lettres de F.

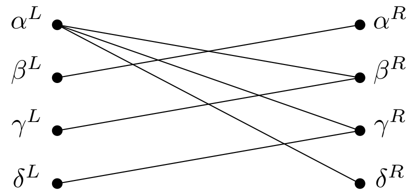
Remarque 2.1.2. Ce morphisme est bien sûr effaçant,de sorte qu'il n'y a a priori pas de lien entre p et $\delta_F(p)$ pour l'inévitabilité.

Définition 2.1.3. Soit p un motif. Le *graphe d'adjacence* de p est défini comme le graphe biparti non-orienté ayant :

- Comme premier ensemble de sommet l'ensemble E^L, similaire à E, dont les élements sont notés $\xi^L, \xi \in E$.
- Comme second ensemble E^R, le même avec des R.
- Comme ensemble d'arêtes l'ensemble des (ξ^L, ζ^R) avec $\xi\zeta$ facteur de p.

On note ce graphe $AG(p)$.

Exemple 2.1.4. Le graphe du motif $\alpha\beta\alpha\gamma\beta\alpha\delta\gamma$ est



Définition 2.1.5. Soit $F \subseteq E$. F est un *ensemble libre* pour p si $\forall(\xi, \zeta) \in F^2$, il n'y a pas de chemin entre ξ^L et ζ^R dans le graphe d'adjacence de p, et si F est non vide.

Exemple 2.1.6. Dans l'exemple 2.1.4,les ensembles libres sont $\{\alpha\}$ et $\{\beta\}$.

Définition 2.1.7. Soit p un motif et F un ensemble libre pour p. On dit que p se *réduit en une étape à q par F* si $q = \delta_F(p)$. On note cette relation $p \xrightarrow{F} q$. On dit donc que p se *réduit en une étape* à q si $\exists F \subseteq alph(p)$ un ensemble libre pour p : $p \xrightarrow{F} q$, ce qu'on note naturellement $p \to q$. Enfin, on dit que p se *réduit* à q si $p \xrightarrow{*} q$, où $\xrightarrow{*}$ est la clôture transitive de $\to$. Un motif que l'on ne peut pas réduire est dit *irréductible.*

Exemple 2.1.8. Reprenons le motif de l'exemple 2.1.4. Ce motif se réduit au motif vide de la façon suivante : $\alpha\beta\alpha\gamma\beta\alpha\delta\gamma \xrightarrow{\{\alpha\}} \beta\gamma\beta\delta\gamma \xrightarrow{\{\beta\}} \gamma\delta\gamma \xrightarrow{\{\gamma\}} \delta \xrightarrow{\{\delta\}} \epsilon$

Remarquons que cela ne signifie pas que toutes ses réductions se réduisent au mot vide. En effet, toutes ses autres réductions terminent sur un motif non vide irréductible.

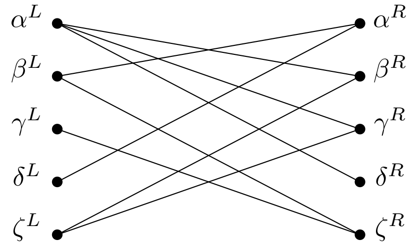
Proposition 2.1.9. *La réductibilité au mot vide est décidable, et est dans NP.*

Démonstration. Algorithme :
Si $p = \epsilon$ accepter.
Prendre un sous-ensemble F de $alph(p)$ de manière non déterministe. (ce qui signifie que il y a un calcul par choix possible pour F)
Si F est libre :
Alors renvoyer le resultat de l'algorithme sur $\delta_F(p)$.
Sinon refuser.

Tester si F est libre se fait évidemment en temps polynomial puisque il s'agit de tester pour les $|F|^2$ éléments de F^2 s'il y a un chemin dans le graphe d'adjacence, qui est calculable en temps polynomial, entre les deux élements du couple, et $|F| \leq |alph(p)| \leq |p|$. Comme la profondeur de récursivité est inférieure à $|p|$, l'algorithme est bien polynomial. (Et il donne évidemment le bon résultat).

$\square$

Remarque 2.1.10. Cet algorithme n'est donc pas très efficace en pratique, surtout que contrairement à ce que laissait penser l'exemple 2.1.8, les réductions menant au mot vide ne passent pas forcément toutes par des singletons. En effet, considérons le motif $\alpha\beta\alpha\gamma\zeta\beta\alpha\delta\alpha\beta\zeta\gamma\zeta\beta\zeta$. Pour ce motif, aucune réduction par un singleton ne mène au mot vide. Son graphe d'adjacence est :



Ce qui donne comme ensembles libres les sous-ensembles de $\{\alpha, \zeta\}$ et $\{\beta, \gamma, \delta\}$. Or, toute réduction par un sous-ensemble du second fait apparaître un facteur α^2 ou ζ^2, qui ne pourra jamais être éliminé par les réductions suivantes (aucun ensemble contenant α/ζ ne pourra être libre). Il reste donc à conidérer les réductions par les sous-ensembles de $\{\alpha, \zeta\}$. Les réductions par $\{\alpha\}$ et $\{\zeta\}$ mènent aux motifs $\beta\gamma\zeta\beta\delta\beta\zeta\gamma\zeta\beta\zeta$ et $\alpha\beta\alpha\gamma\beta\alpha\delta\alpha\beta\gamma\beta$, dont les graphes d'adjacence sont connexes, et donc sont irréductibles. Or on a la réduction : $\alpha\beta\alpha\gamma\zeta\beta\alpha\delta\alpha\beta\zeta\gamma\zeta\beta\zeta \xrightarrow{\{\alpha,\zeta\}} \beta\gamma\beta\delta\beta\gamma\beta \xrightarrow{\{\beta\}} \gamma\delta\gamma \xrightarrow{\{\gamma\}} \delta \xrightarrow{\{\delta\}} \epsilon$.

On a alors le théorème suivant :

Théorème 2.1.11. *Un motif est inévitable si et seulement si il est réductible au mot vide.*

Corollaire 2.1.12. *Si un motif p est inévitable alors il existe une variable de p qui n'apparaît qu'une fois.*

Démonstration. Si toute variable de p apparaît au moins deux fois, $p \neq \epsilon$, et si $p \xrightarrow{F} q$, alors $F \neq alph(p)$. En effet, $AG(p)$ contient un arc car $|p| \geq 2$, donc $F \neq alph(p)$. En particulier, toute variable de q apparaît au moins deux fois et $q \neq \epsilon$. Une récurrence immédiate donne $\forall q$ motif, $p \xrightarrow{*} q \Rightarrow q \neq \epsilon$. Donc p n'est pas réductible au mot vide et donc évitable.

$\square$

Corollaire 2.1.13. *Si p est un motif inévitable à n variables, alors $|p| < 2^n$.*

Démonstration. Par récurrence : c'est trivial si $n = 0$. Si p est un motif à $n + 1$ variables, et que la propriété est vraie pour n : Par le corollaire précédent il existe une variable α dans p qui n'apparaît qu'une fois : $p = p_1\alpha p_2$ avec p_1 et p_2 inévitables à n variables : l'inégalité est alors immédiatement vérifiée par hypothèse de récurrence.

$\square$

La preuve se découpe en deux parties, une par sens.

2.2 Les motifs réductible au mot vide sont inévitable

Proposition 2.2.1. *Soit p,q deux motifs tels que* $p \xrightarrow{F} q$*. Alors si q est inévitable, p aussi.*

Corollaire 2.2.2. *Par récurrence immédiate, si p,q sont deux motifs tels que* $p \xrightarrow{*} q$*, si q est inévitable alors p aussi. En particulier, si* $q = \epsilon$*, p est inévitable.*

Démonstration. On va procéder par récurrence sur la taille de l'alphabet A. Le cas $|A| = 1$ est immédiat car p est inévitable sur A sans aucune autre hypothèse. Supposons donc que $|A| = n + 1$ et que p soit inévitable sur les alphabets de taille n , on fixe alors a une lettre de A : $A = A' \cup \{a\}$, et p est inévitable sur A. Donc en particulier l'ensemble $L = A'^{+} \backslash A'^{*} p(A'^{+}) A'^{*}$ est fini. On pose $M = aA^{*} \backslash A^{*} p(A^{+}) A^{*}$. Tout élément de M qui n'est pas une puissance de a peut donc s'écrire comme produit d'éléments de $N = \{a^{i} \omega a^{j} | 0 < i < |p|, 0 \leq j < |p|, \omega \in L\}$, car les occurrences de mots de A' de longueur maximale dans un mot de M sont forcément dans L, et sont séparées des puissances de a d'ordre inférieur à p car les mots de M évitent p. Comme N est fini, on peut voir N comme un nouvel alphabet, et on alors un morphisme naturel $i : N^{*} \to A^{*}$ obtenu en remplaçant une lettre de N par elle-même vue comme un mot de A^{*}. On a alors $M \subseteq i(N^{+}) \cup a^{+}$. Il suffit donc de montrer que l'ensemble des mots de $i(N^{+})$ qui évitent p est fini : en effet, l'ensemble des mots de M qui sont des puissances de a est fini car elles évitent p, et les autres élements de M sont dans $i(N^{+})$ et évitent p. L'ensemble des mots de A^{*} qui évitent p et qui commence par a serait alors fini, et il en serait donc de même pour chaque lettre de A par permutation. Donc si l'ensemble des mots de $i(N^{+})$ qui évitent p est fini, l'ensemble des mots de A^{*} qui évitent p est fini également.
On introduit ζ une nouvelle variable ($\xi \notin alph(p)$). Le motif $q\xi$ est évidemment inévitable, puisque par proposition 1.3.1, $q\xi q$ est inévitable. Donc $\exists N_0 \in \mathbb{N}$:
$\forall \omega \in N^{*}, |\omega| \geq N_0 \Rightarrow \omega$ n'évite pas p. Soit $l_0 = max\{|\omega| | \omega \in N\}$ (vus comme des mots). Soit $\omega \in i(N^{+})$ quelconque tel que $|\omega| \geq l_0 N_0$: $\exists u \in N^{+} : \omega = i(u)$, et évidemment $|u| \geq N_0$. Donc il existe un morphisme non-effaçant f de $(E \cup \{\xi\})^{*}$ dans N^{*} tel que $f(q\xi)$ soit facteur de u. Il suffit maintenant de construire un morphisme non-effaçant g de E^{*} dans A^{*} tel que $g(p)$ soit facteur de $i(f(q\xi))$, puisque $i(f(q\xi))$ est facteur de ω. On va noter C^{R} l'ensemble des élements de E^{R} dans $AG(p)$ qui sont reliés par un chemin a F^{L}, et C^{L} le même avec des L. Soit donc g le morphisme défini par :

$$g(\delta) = \begin{cases} a & \text{si } \delta \in F \\ i(f(\delta)) & \text{si } \delta^{R} \notin C^{R} \text{ et } \delta^{L} \notin C^{L} \\ a^{-1} i(f(\delta)) & \text{si } \delta^{R} \in C^{R} \text{et } \delta^{L} \notin C^{L} \\ i(f(\delta)) a & \text{si } \delta^{R} \notin C^{R} \text{ et } \delta^{L} \in C^{L} \text{ et } \delta \notin F \\ a^{-1} i(f(\delta)) a & \text{si } \delta^{R} \in C^{R} \text{ et } \delta^{L} \in C^{L} \end{cases}$$

Le principe est qu'on cherche à effacer la contribution de F, tout en obtenant $f(q\xi)$. Pour cela, les lettres dans F sont quasi-effacées (remplacées par une seule lettre), et pour effacer ces a, on enlève a au début de $i(\delta)$ quand $\delta^{R} \in C^{R}$. Or, cela va enlever également des a à la fin des images des δ avec $\delta^{L} \in C^{L}$, donc pour ceux-là (sauf ceux de F évidemment) on rajoute un a à la fin.
Ce morphisme est bien défini parce que F est un ensemble libre : les quatre dernières conditions sont trivialement disjointe, et les élements qui ne vérifient aucune de ces conditions sont les élements de F : ceux-ci sont forcément dans C^{L}, et ne peuvent être dans C^{R} par liberté.
Montrons maintenant que pour tout préfixe p' de p, $r_{p'} g(p') = i(f(\delta_F(p'))) s_{p'}$ où $r_{p'}$ vaut a si la première lettre de p' est dans C^{R} et ϵ sinon, $s_{p'}$ vaut a si la dernière lettre de p' est dans C^{L} et ϵ sinon. On procède par récurrence sur $|p'|$: pour $|p'| = 1$ c'est la définition de g ($F \subseteq C^{L}$). Placons-nous maintenant dans le cas où

$p' = p''\beta$, et où la propriété est vraie pour p'' :
$r_{p'}g(p') = r_{p''}g(p''\beta) = r_{p''}g(p'')g(\beta) = i(f(\delta_F(p''))s_{p''}g(\beta)$
Or, si la dernière lettre γ de p'' est dans C^L, $\beta \in C^R$ par définition de C^R et C^L (il y a une arête entre γ^L et β^R). Donc $s_{p''}g(\beta) = i(f(\beta))s_{p'}$ si $\beta \notin F$, et sinon $s_{p''} = \epsilon$ car $\beta \notin C^R$, et $g(\beta) = a = i(f(\epsilon))s_{p'}$, soit $i(f(\delta_F(\beta)))s_{p'}$ dans tout les cas. En utilisant le fait que i, f, et δ_F sont des morphismes on obtient l'hérédité. Donc en particulier pour p : $r_p g(p) = i(f(q))s_p$. C'est là qu'on voit pourquoi on a rajouté une variable à q : pour faire rentrer le s_p dans le mot final. En effet, a est préfixe de $i(f(\xi))$, donc s_p en est préfixe. Donc $r_p g(p)$ est facteur de $i(f(q\xi))$, donc $g(p)$ en est facteur. C'est ce qu'il restait à démontrer.

□

2.3 Les motifs inévitables sont réductibles au mot vide

Lemme 2.3.1. *Soient p, q deux motifs et f un morphisme non-effaçant tels que $f(q)$ est facteur de p (donc $q|p$). Soit F un ensemble libre pour p. Alors $F' = \{\xi \in F | f(\xi) \in F^+\}$ est un ensemble libre pour q, et si : $p \xrightarrow{F} p'$, $q \xrightarrow{F'} q'$, alors $\exists f'$ morphisme non-effaçant de $(alph(q)\backslash F')^*$ dans $(alph(p)\backslash F)^*$ tel que $f'(q')$ soit facteur de p'.*

Démonstration. Soit g de $alph(q)^*$ dans $alph(p)^*$ défini en prenant la première lettre de f, et h de même la dernière. Soient $(\xi, \zeta) \in alph(q)^2$: Si il y a une arête entre ξ^L et ζ^R dans $AG(q)$: $\xi^L\zeta^R$ est facteur de q, et donc $h(\xi)g(\zeta)$ est facteur de p : il y a donc une arête entre $h(\xi)^L$ et $g(\zeta)^R$ dans $AG(p)$. En envoyant β^L sur $h(\beta)^L$ et β^R sur $g(\beta)^R$, on obtient donc que $AG(q)$ est envoyé sur un sous-graphe de $AG(p)$. Soient donc $(\xi, \zeta) \in F'^2$ quelconques : Si il y a un chemin entre ξ^L et ζ^R dans AG(q), il y en a un entre $h(\xi)^L$ et $g(\zeta)^R$, ce qui est absurde car F est libre. Donc F' est libre.
Posons maintenant $f' = \delta_F \circ f_{|(alph(q)\backslash F')^*}$. Si $f(q)$ est facteur de p, il est immédiat que $\delta_F(f(q))$ est facteur de $p' = \delta_F(p)$. Or, $\forall \xi \in F'$, $\delta_F(f(\xi)) = \epsilon$. Donc
$\delta_F(f(q)) = \delta_F(f(\delta_{F'}(q))) = \delta_F(f(q')) = f'(q')$, donc $f'(q')$ est bien facteur de p'. De plus, f' est bien non-effaçant car $\forall \xi \in alph(q)$, $\delta_F(f(\xi)) = \epsilon \Leftrightarrow \xi \in F'$, et f' n'est pas défini sur F'.

□

Lemme 2.3.2. *Si p, q sont deux motifs et f un morphisme non-effaçant tel que $f(q)$ est facteur de p, et si $p \xrightarrow{*} p'$, il existe un mot q' et un morphisme non-effaçant f' tels que $q \xrightarrow{*} q'$, $f'(q')$ est facteur de p', $f((alph(q)\backslash alph(q'))) \subseteq (alph(p)\backslash alph(p'))^*$, et que de plus $f' = \delta_{alph(p)\backslash alph(p')} \circ f_{|alph(q')^*}$.*

Démonstration. La preuve se fait par itération du lemme 2.3.1. Il suffit de le vérifier pour $p = p'$ et pour $p \xrightarrow{*} p'' \rightarrow Fp'$ avec la propriété vérifié par p et p'' (récurrence descendante sur la longueur de p'). Pour $p = p'$, c'est immédiat avec $f' = f$ et $q' = q$. Plaçons-nous donc dans la seconde situation. Il existe donc q'' et f'' satisfaisant la propiété demandée pour p''. On applique alors le lemme 2.3.1 : soient q' le mot obtenu et f' le morphisme obtenu, F' l'ensemble libre donné par ce lemme. Alors $f'(q')$ est facteur de p'. Or la construction fournie par ce lemme donne $f''(F') \subseteq F^*$, donc $f(F') \subseteq (F \cup (alph(p)\backslash alph(p'')))^* = (alph(p)\backslash alph(p'))^*$.

$$\begin{aligned} f(alph(q)\backslash alph(q')) &= f((alph(q)\backslash alph(q'')) \cup F') \\ &= f(alph(q)\backslash alph(q'')) \cup f(F') \\ &\subseteq ((alph(p)\backslash alph(p''))^*) \cup (alph(p)\backslash alph(p'))^* \\ &\subseteq (alph(p)\backslash alph(p'))^* \end{aligned}$$

Il reste à vérifier la formule donnant f' :

$$\begin{aligned} f' &= \delta_F \circ f''_{|alph(q')^*} \\ &= \delta_F \circ \delta_{alph(p)\backslash alph(p'')} \circ (f_{|alph(q'')^*})_{|alph(q')^*} \\ &= \delta_{alph(p)\backslash alph(p')} \circ f_{|alph(q')^*} \end{aligned}$$

Ce qui conclut le lemme.

$\square$

Lemme 2.3.3. *Soit p un motif et V un ensemble de variables de p. Soit $q = \delta_V(p)$. Supposons qu'il existe un motif r et un morphisme non-effaçant f tels que $r \xrightarrow{*} q$, $f(p)$ soit facteur de r, et $\forall \zeta \in alph(p)$, $f(\zeta) \in (alph(r)\backslash alph(q))^* \Rightarrow \zeta \in V$. Alors $p \xrightarrow{*} q$.*

Démonstration. Appliquons le lemme 2.3.2 à p, r et f : $\exists q' \in alph(p)^*$, f' morphisme, non-effaçant de $alph(p')^*$ dans $alph(q)^*$: $f'(p')$ est facteur de q, $p \xrightarrow{*} p'$ et $f(alph(p)\backslash alph(p')) \subseteq (alph(r)\backslash alph(q))^*$, donc $alph(p)\backslash alph(p') \subseteq V$. Donc $\delta_V(p') = \delta_V(p) = q$. Or cela donne que $|p'| \geq |q|$, et on a l'inégalité inverse immédiatement par $f'(p')$ facteur de q. Donc $p' = q$, ce qui conclut le lemme.

$\square$

Définition 2.3.4. Pour tout $k \in \mathbb{N}$, on pose φ_k le morphisme défini sur l'ensemble $\{a_0, a_1, ..., a_{2k-1}, b_0, b_1, ..., b_{2k-1}\}^*$ par :
- $\forall i \in [|0, 2k-1|], \varphi_k(a_i) = a_0 b_i a_1 b_{i+1} ... a_{k-1} b_{i+k-1}$
- $\forall i \in [|0, 2k-1|], \varphi_k(b_i) = a_k b_i a_{k+1} b_{i+1} ... a_{2k-1} b_{i+k-1}$.

Les indices sont pris dans $\mathbb{Z}/2k\mathbb{Z}$. Par lemme 1.2.1, ce morphisme admet un unique point fixe infini commençant par a_0, que l'on appelle $\omega^{(k)}$.
Notons par ailleurs que φ_k multiplie la longueur de tout mot par $2k$: $\omega^(k)$ est donc composé de blocs de taille $2k$ qui sont les images par φ_k des lettres de lui-même.

Lemme 2.3.5. *Soit v un facteur de longueur 2 de $\omega^{(k)}$. Alors $\exists i_v \in \mathbb{Z}/4k\mathbb{Z}$ et $x_v \in alph(\omega^{(k)})$: $\forall n \in \mathbb{N}$, si v est facteur en position n :*

1. *(i) $n \equiv i_v[4k]$*
2. *(ii) $\omega^{(k)}_{\lfloor \frac{n}{2k} \rfloor} = x_v$*

Ces conditions exprime que à partir du facteur v, il est possible de retrouver la lettre de $alph(\omega^{(k)})$ donnant comme image le bloc de taille $2k$ dans lequel se trouve la première lettre de v, ainsi que la position de ce facteur dans $\omega^{(k)}$ modulo $4k$. On dit alors que un tel facteur v reconnaît $\varphi_k(x)$.

Démonstration. Remarquons d'abord que les blocs commençant par a_0 et a_k alternent dans $\omega^{(k)}$ puisque les a_i et les b_j alternent dans celui-ci. Donc les lettres de $\omega^{(k)}$ à des positions paires sont $a_0, a_1, ..., a_{2k-1}$ dans l'ordre pour les $2k-1$ premières, et sont $2k$-périodiques. Puisque v contient une lettre a_i, il suffit de regarder cet indice pour obtenir i_v à un près : la lettre a_i apparaît forcément à une position congrue à $2i$ modulo $4k$ (les lettres en position impaires sont des b_j). Pour obtenir i_v exactement il faut regarder laquelle des deux lettres de v est a_i.
Remarquons maintenant que l'on sait trouver explicitement $x_n = \omega^{(k)}_{\lfloor \frac{n}{2k} \rfloor}$, si v apparît en position n : en effet la premiére lettre de v est dans le bloc $\varphi_k(x_n)$, en combinant avec le fait que la lettre deux positions a_i est a_{i+1} on obtient :
- si $v = a_i b_j$ avec $i < k$: nécessairement $x_n = a_{j-i}$
- si $v = a_{i+k} b_j$ avec $i < k$: $x_n = b_{j-i}$
- si $v = b_j a_{i+1}$ avec $i < k$: $x_n = a_{j-(i-1)}$
- si $v = b_j a_{i+k+1}$ avec $i < k$: $x_n = b_{j-(i-1)}$.

Comme cela ne dépend évidemment pas de n on obtient le x_v cherché.

$\square$

Proposition 2.3.6. *Soit p un motif, $k \in \mathbb{N}$ tel que $2k > card(alph(p))$, et v un facteur de $\omega^{(k)}$ tel que $\varphi(v)$ n'évite pas p : alors il existe un motif q tel que $p \xrightarrow{*} q$ et que v n'évite pas q.*

Démonstration. Puisque $\varphi_k(v)$ n'évite pas p, il existe μ un morphisme non-effaçant de $E = alph(p)$ dans $A = alph(\omega^{(k)})$ tel que $\mu(p)$ soit facteur de $\varphi_k(v)$. Par lemme 2.3.5, pour tout $x \in A$, il existe précisément $2k$ mots reconnaissant $\varphi_k(x)$. Or ces mots se terminent tous par des lettres différentes : il en existe donc un, que l'on note d_x et que l'on appelle le *mot décisif* pour x, qui termine par une lettre qui n'est la première lettre d'aucun $\mu(\zeta)$ pour $\zeta \in E$, puisque $2k > |E|$. La construction de d_x permet d'affirmer que si d_x est facteur de $\mu(p)$ alors il existe une lettre $\zeta \in E$ telle que d_x soit facteur de $\mu(\zeta)$.
Soit V l'ensemble des variables ζ telles que $\mu(\zeta)$ ne contienne pas de mot décisif, et $q = \delta_V(p)$. Posons maintenant ν le morphisme non effaçant de $(E\backslash V)^*$ dans A^* suivant : $\nu(\zeta) = x_1x_2...x_m$ où $d_{x_1}, d_{x_2}, .., d_{x_m}$ sont les mots décisifs de $\mu(\zeta)$ dans leur ordre d'apparition. Il est alors immédiat que $\nu(q)$ est facteur de v : En effet $\nu(q) = x_1...x_m$ où les $d_{x_1}, d_{x_2}, .., d_{x_m}$ sont les mots décisifs facteurs de $\mu(p)$ dans leur ordre d'apparition. Or si $d_{x'_1}, d_{x'_2}, ...d_{x'_{m'}}$ sont les mots décisifs de $\varphi_k(v)$, le lemme 2.3.5 nous assure que $v = x_1..x_m$, et le fait que $\mu(p)$ soit facteur de $\varphi_k(v)$ assure que les d_{x_i} sont de la forme :$d_{x_i} = d_{x'_{i+j}}$ pour un certain j, d'où $\nu(q)$ est bien facteur de v. Cependant, il n'y a a priori aucune raison pour que V soit un ensemble libre, et donc que $p \xrightarrow{*} q$, ce qui est nécessaire pour conclure.
On va donc appliquer le lemme 2.3.3. Il suffit pour cela de construire un morphisme f adéquat. On va pour le construire supposer E et A disjoint (si ce n'est pas le cas on les rends artificiellement disjoints). On va poser un morphisme f de E^* dans $(E \cup A)^*$. Pour qu'il fonctionne, on va avoir besoin que $f(p)$ se ramène à q par délétion de variables. Comme on veut en plus que ces variables soient dans un ensemble libre, on va essayer d'avoir $q = \delta_A(f(p))$ puisque A contient les variables qu'il est le plus facile de contrôler. Enfin, comme les facteurs de $\omega^{(k)}$ se prêtent bien à la réduction, on va essayer d'avoir de tels facteurs entre les lettres de q. Enfin il faut bien sûr qu'il satisfasse les hypothèses du lemme. Le morphisme est le suivant :

- Si $\zeta \in V$: $f(\zeta) = \mu(\zeta)$.
- Si $\zeta \notin V$ et $\mu(\zeta)$ ne contient ni a_0 ni a_k : $\mu(\zeta)$ contient précisément un mot décisif d_x , et on a : $\varphi_k(x) = a_iv_1\mu(\zeta)v_2$ pour $i \in \{0, k\}$, et v_1, v_2 deux mots de A^*. On pose $f(\zeta) = \mu(\zeta)v_2\zeta v_1\mu(\zeta)$
- dernier cas : on a donc $\mu(\zeta) = v_1\varphi_k(u)a_jv_2$ avec v_1, v_2, u trois mots de A^*, u de longueur maximale (ni a_0 ni a_k n'apparaissent dans v_1 et v_2) et $j \in \{0, k\}$. Soit a_i la première lettre de $\varphi_k(u)$. On pose alors : $f(\zeta) = v_1v'_1\zeta v'_2v_2$ avec $v'_1 = \epsilon$ si le premier mot décisif d_{x_1} de $\mu(\zeta)$ est dans le facteur v_1a_i, $v'_1 = \varphi_k(x_1)$ sinon, et de même $v'_2 = \epsilon$ si le dernier mot décisif d_{x_2} de $\mu(\zeta)$ est dans le facteur a_jv_2, $(a_{j+k})^{-1}(\varphi(x_2))a_j$ sinon.

Le morphisme qu'on a ici vérifie bien la propriété souhaitée pour la délétion de variable de A, ainsi que la propriété du lemme. De plus, les propriétés complexes qui définissent le morphisme f assurent que le facteur entre de deux lettres consécutives ζ_1 et ζ_2 de q dans $f(p)$ est $v_1a_iv_2$, où $a_{i+k}v_1$ est l'image de la première lettre de $\nu(\zeta_1)$ par φ_k et a_iv_2 l'image de la dernière lettre de $\nu(\zeta_2)$ par φ_k. En effet, si ζ_1, ζ_2 sont deux lettres consécutives de q séparées par $p_1, p_2, ..p_m$ dans p (m peut être nul) : $f(\zeta_1p_1...p_m\zeta_2)$ $= (...)\zeta_1u_1\mu(p_1...p_m)u_2\zeta_2(...)$. Or les propriétés de f montrent que soit (deuxième cas) $\mu(\zeta_1)$ est un suffixe de u_1 : $u_1 = v_1\mu(\zeta_1)$, et la définition de f assure que les lettres de v_1 sont bien celles qui précèdent $\mu(\zeta_1)$ dans $\omega^{(k)}$ par le lemme 2.3.5, soit (troisième cas) u_1 est un suffixe de $\mu(\zeta_1)$. Avec le même genre de raisonnement pour ζ_2, on obtient que le facteur u_0 de $f(p)$ entre ζ_1 et ζ_2 est bien un facteur de $\omega^{(k)}$, commençant évidemment une lettre après la première lettre d'un bloc de taille $2k$ de $\omega^{(k)}$ (il suit a_i, $i \in \{0, k\}$) et finissant de même au niveau de la fin d'un bloc de taille 2k de $\omega^{(k)}$. Or, la définition de f assure que a_iu_1 et u_2 contiennent chacun exactement un mot décisif, et la définition de ces mots assure donc que a_iu_0 contient exactement deux mots décisif, qui sont le dernier de $\mu(\zeta_1)$ et le premier de $\mu(\zeta_2)$, d'où la propriété annoncée sur le facteur.
Maintenant qu'on a ce morphisme,il faut montrer que $f(p) \xrightarrow{*} q$ pour terminer la preuve. Commençons par remarquer que $B = \{b_i | i \in [|0, 2k-1|]\}$ est trivialement

libre,puisque ces lettres ne côtoient que des a_j dans $f(p)$. Donc $f(p) \xrightarrow{*} \delta_B(f(p))$. Or, remarquons maintenant que les facteurs entre deux lettres consécutives de q dans $\delta_B(f(p))$ ne peuvent être que $a_1a_2...a_{2k-1}$ ou $a_{k+1}a_{k+2}...a_{k-1}$. Ces lettres peuvent également apparaître dans $\delta_B(f(p))$ avant toute les lettres de q ou après toutes ces lettres mais un raisonnement similaire à celui utilisé pour trouver la forme des facteurs intérmédiaire montre que ce sont nécessairement des préfixes ou suffixes de ces mots. Remarquons alors que $\{a_1, a_{k+1}\}$ est un ensemble libre. En effet, a_1^L n'est connecté qu'à a_2^R et vice-versa, tandis que a_1^R est connecté à a_0^L, lequel n'est connecté qu'à celui-ci, et à des lettres de la forme ζ^R, $\zeta \in E\backslash V$. Or la définition de f montre directement que ces sommets ne sont connectés qu'à a_1^R aussi. Donc en particulier il n'y a pas de chemin entre a_1^L et a_{k+1}^R, ni entre a_1^R et a_{k+1}^L. C'est donc bien un ensemble libre, noté F_1. Donc $f(p) \xrightarrow{*} p_1 = \delta_{B\cup F_1}(f(p))$: Or le même argument peut alors être employé pour $\{a_2, a_{k+2}\}$. On retire alors successivement tout les $F_i = \{a_i, a_{k+i}\}$ pour $i \in [|1, k-1|]$, et on obtient un motif p_{k-1} composé des lettres de q séparées par des a_0 et des a_k, avec peut-être une de ces lettres en préfixe/suffixe. Or il est immédiat que $\{a_0\}$ est un ensemble libre, puis que $\{a_k\}$ est libre dans le motif résultant. Donc $f(p) \xrightarrow{*} q$, ce qui clôt la preuve.

□

Corollaire 2.3.7. *Soit p un motif, $k \in \mathbb{N}$ tel que $2k > |alph(p)|$. Si $\omega^{(k)}$ n'évite pas p alors p se réduit au mot vide (et est donc inévitable). En particulier, tout motif inévitable se réduit au mot vide, et $\omega^{(k)}$ évite tout les motifs évitables avec au plus $2k-1$ variables. Cela termine d'ailleurs la démonstration du théorème 2.1.11.*

Démonstration. Suposons que $\omega^{(k)}$ n'évite pas p. Alors il existe un préfixe fini v de $\omega^{(k)}$ qui n'évite pas p. Donc il existe $m \in \mathbb{N}$ tel que v soit préfixe de $\varphi_k^m(a_0)$ d'aprés la construction de $\omega^{(k)}$ donnée par le lemme 1.2.1. Or en appliquant la proposition 2.3.6 itérativement m fois, on obtient qu'il existe un motif q : $p \xrightarrow{*} q$ et a_0 n'évite pas q, ce qui n'est possible que si q n'a au plus qu'une lettre. Donc $q \xrightarrow{*} \epsilon$, et $p \xrightarrow{*} \epsilon$.

□