Annexe A

Classification partielle des motifs ternaires

Une fois la classification des motifs binaires achevée, nous avons entrepris d'en faire autant avec les motifs ternaires. Silke Nilgens [47] avait déjà montré que presque tous les motifs ternaires évitables sont 3-évitables. En utilisant les techniques développées dans cette thèse, et en particulier l'algorithme présenté au chapitre 6, ainsi que quelques résultats existants, nous avons pu compléter ses résultats et statuer sur la 2-évitabilité d'un grand nombre de motifs. Comme pour les motifs binaires, nous avons aussi essayé de déterminer ceux des motifs ternaires 2-évitables qui sont 2-DOL-inévitables, et ceux des motifs 3-évitables qui sont 3-DOL-inévitables.

Nous présentons ici l'état actuel de cette classification. À permutation des lettres près, il reste à classer : 1 motif pour la 3-évitabilité (de longueur 7), 5 motifs pour la 3-DOL-évitabilité (de longueur 6 et 7), 103 motifs pour la 2-évitabilité (de longueur 6 à 10), et 193 motifs pour la 2-DOL-évitabilité (de longueur 7 à 13), soit en tout 274 motifs (certains sont dans plusieurs listes).

On constate que la longueur des motifs 2-inévitables connus ne dépasse pas 11, celle des motifs 2-DOL-inévitables ne dépasse pas 14, et celle des motifs 3-DOL-inévitables ne dépasse pas 7. Comme aucun motif inconnu n'est de longueur supérieure, ces trois valeurs sont bien les longueurs maximales.

A.1 Les résultats de Silke Nilgens

Pour prouver la 3-évitabilité de certains motifs ternaires, S. Nilgens utilise le morphisme suivant, sur $\Sigma = \{a, b, c\}$:

$$\begin{array}{rccl} \delta : & \Sigma^* & \longrightarrow & \Sigma^* \\ & a & \longmapsto & ab \\ & b & \longmapsto & bc \\ & c & \longmapsto & ca \end{array} .$$

Le mot infini qu'il engendre, $\delta^\omega(a)$, peut être vu comme une généralisation de la suite de Thue-Morse; comme ce dernier, c'est un mot sans chevauchement. Le fait que le morphisme δ soit uniforme facilite les démonstrations d'évitabilité, puisqu'on peut utiliser des arguments de parité.

S. Nilgens montre que $\delta^\omega(a)$ évite les motifs $ABCABAC$, $ABCBABCA$, $ABACBCA$ et $ABCACBA$. Comme on sait par ailleurs que AA est 3-évitable, elle peut donc conclure que tous les motifs ternaires de longueur supérieure ou égale à 8, qui sont divisibles par l'un des motifs ci-dessus (ou par leur image miroir), sont 3-évitables. La borne ainsi obtenue est optimale, car $ABACABA$, de longueur 7, est inévitable. En fait, la plupart des motifs restant sont 3-inévitables; seuls 5 motifs, à permutation des lettres et image miroir près, échappent à cet argument : $ABACBC$, $ABCACB$, $ABACABC$, $ABACBAB$ et $ABCBABC$.

A.2 Les 5 motifs restants

Tous les motifs dont la 3-évitabilité a été établie par S. Nilgens sont en fait 3-DOL-évitables. Pour les motifs restants, nous allons devoir utiliser des HDOL-langages, car $ABCBABC$ et $ABACBAB$ sont 3-DOL-inévitables et il n'a pas été possible de trouver de DOL-langage pour les trois autres.

Soient $\Sigma_4 = \{a, b, c, d\}$ et $\Sigma_3 = \{a, b, c\}$, et les morphismes suivants :

$$\begin{array}{rrcl} f\colon & \Sigma_4^* & \longrightarrow & \Sigma_4^* \quad , \\ & a & \longmapsto & ab \\ & b & \longmapsto & ac \\ & c & \longmapsto & dc \\ & d & \longmapsto & db \end{array}$$

$$\begin{array}{rrcl} g_1\colon & \Sigma_4^* & \longrightarrow & \Sigma_3^* \quad , \\ & a & \longmapsto & ab \\ & b & \longmapsto & bc \\ & c & \longmapsto & aa \\ & d & \longmapsto & cb \end{array} \qquad \begin{array}{rrcl} g_2\colon & \Sigma_4^* & \longrightarrow & \Sigma_3^* \quad . \\ & a & \longmapsto & ab \\ & b & \longmapsto & ac \\ & c & \longmapsto & bc \\ & d & \longmapsto & cb \end{array}$$

Au moyen de l'algorithme présenté au chapitre 6, on montre que le HDOL-langage $L(\Sigma_4, f, a, \Sigma_3, g_1)$ évite les motifs $ABACBC$ et $ABACABC$, et que le HDOL-langage $L(\Sigma_4, f, a, \Sigma_3, g_2)$ évite les motifs $ABACABC$, $ABACBAB$ et $BCACBA$ (image miroir de $ABCACB$). Quatre des cinq motifs étudiés sont donc 3-HDOL-évitables.

Le dernier motif, $ABCBABC$, pose nettement plus de problèmes. Les résultats du *backtracking* portent à croire qu'il est 3-évitable, et même 2-évitable. Nous pensons qu'il est évité par les HDOL-systèmes $G_1 = (\Sigma_5, f_1, a, \Sigma_3, g_1)$ et

$G_2 = (\Sigma_5, f_2, a, \Sigma_2, g_2)$ où $\Sigma_5 = \{a, b, c, d, e\}$, $\Sigma_3 = \{a, b, c\}$, $\Sigma_2 = \{a, b\}$ et

$$f_1 \colon \begin{array}{ccc} \Sigma_5 & \longrightarrow & \Sigma_5 \\ a & \longmapsto & abad \\ b & \longmapsto & cbce \\ c & \longmapsto & bda \\ d & \longmapsto & bcd \\ e & \longmapsto & edbde \end{array} , \qquad g_1 \colon \begin{array}{ccc} \Sigma_5 & \longrightarrow & \Sigma_3 \\ a & \longmapsto & aa \\ b & \longmapsto & ccca \\ c & \longmapsto & acbb \\ d & \longmapsto & cab \\ e & \longmapsto & accb \end{array} ,$$

$$f_2 \colon \begin{array}{ccc} \Sigma_5 & \longrightarrow & \Sigma_5 \\ a & \longmapsto & ab \\ b & \longmapsto & ac \\ c & \longmapsto & ad \\ d & \longmapsto & ce \\ e & \longmapsto & be \end{array} , \qquad g_2 \colon \begin{array}{ccc} \Sigma_5 & \longrightarrow & \Sigma_2 \\ a & \longmapsto & aaab \\ b & \longmapsto & bbba \\ c & \longmapsto & babb \\ d & \longmapsto & abaa \\ e & \longmapsto & baaa \end{array} .$$

Le premier a été obtenu en essayant des morphismes aléatoires; le second, bien plus agréable car uniforme et sur l'alphabet binaire, a été découvert par Kirby Baker. En théorie, il suffirait d'appliquer notre algorithme pour vérifier ces conjectures; mais s'il marche bien avec 2 ou 3 lettres, l'explosion combinatoire fait que les implémentations actuelles n'ont pas de ressources suffisantes pour traiter un alphabet à 5 lettres.

A.3 Résultats de 2-inévitabilité

Par *backtracking*, on arrive à montrer qu'un certain nombre de motifs ternaires sont 2-inévitables. Il peut s'avérer très intéressant de transformer auparavant le motif en une formule équivalente. Ainsi, pour le motif $AABCBABA$, on trouve 224924033 mots qui évitent ce motif, les plus longs étant de longueur 84, par exemple :

$$abababababababaaaaababaaababbaababbaabbbaabbbaaabbbaaabbbaabbaaabbbaabbabbbaaabbbabb$$

(plusieurs jours de calcul sur SPARCstation 10 ont été nécessaires pour terminer le *backtracking*). Par contre, avec la formule équivalente $AAB.BABA$, on ne trouve plus que 719 mots qui l'évitent, de longueur au plus 18, par exemple $abaabbaaabbbaabbab$, et le calcul dure quelques fractions de seconde.

On montre ainsi que les motifs suivants sont 2-inévitables :
$AABABCBAA$, $AABACABB$, $AABACBB$, $AABACBBA$, $AABBCBAA$, $AABBCBBA$, $AABCABBA$, $AABCBABA$, $AABCCAC$, $AABCCB$, $ABAACBAAB$, $ABACABBA$, $ABACCB$, $ABBACAAB$, $ABBCCA$, $ABCBCA$, $ABCCAB$, $ABCCBA$.

A.4 Résultats de DOL-inévitabilité

Les motifs suivants sont 2-DOL-inévitables car simultanément 2-inévitables avec AAA et divisant le multicube C_3 (équivalent à $(A(BCCCD)^3E)^3$, voir page 59) :
$AABAACABAABA$, $AABAACBAAB$, $AABABCBABA$, $AABABCC$, $AABACAABB$, $AABACAABBA$, $AABACABBA$, $AABACBABA$, $AABACBC$, $AABBACAABA$, $AABBACABA$, $AABBACC$, $AABBCAABA$, $AABBCABA$, $AABBCBAB$, $AABBCBC$, $AABBCCB$, $AABCACAB$, $AABCACB$, $AABCBAC$, $AABCBBAB$, $AABCBBC$, $AABCCAB$, $AABCCBA$, $AABCCBC$, $ABAACABAB$, $ABAACABBA$, $ABAACBBA$, $ABAACCAB$, $ABABBCBABA$, $ABABBCBBA$, $ABABBCCB$, $ABABCAAC$, $ABABCAC$, $ABABCBABA$, $ABABCBBA$, $ABABCCB$, $ABACAABBA$, $ABACAACB$, $ABACABC$, $ABACACB$, $ABACACBA$, $ABACBABC$, $ABACBAC$, $ABACBBC$, $ABACBCA$, $ABACCAAB$, $ABACCAB$, $ABACCABA$, $ABBACCA$, $ABBCABC$, $ABBCBAC$, $ABBCBCA$, $ABBCBCAB$, $ABBCCAB$, $ABBCCBA$, $ABCAACB$, $ABCABABC$, $ABCACAB$, $ABCACABC$, $ABCACBA$, $ABCACCAB$, $ABCBBAC$, $ABCBBCBA$.

Ceux-ci sont 2-DOL-inévitables, d'après le théorème 3.4, car contenus dans la suite de Thue-Morse et divisant le multichevauchement O_3.
$AABABCAABCA$, $AABABCACAA$, $AABACABACCA$, $AABACBCA$, $AABACBCC$, $AABACCABAA$, $AABACCABACA$, $AABACCBA$, $AABBCAABC$, $AABBCABBCB$, $AABBCABCA$, $AABBCACA$, $AABCAABABCAAB$, $AABCABAC$, $AABCACABCAC$, $AABCCBAA$, $AABCCBAC$, $AABCCBCA$, $ABAABACABAC$, $ABAABACAC$, $ABAABACBAABAC$, $ABAABACBAC$, $ABAABACCA$, $ABAABCCA$, $ABAACAABAACA$, $ABAACACBAAC$, $ABAACBAACAC$, $ABAACBABAAC$, $ABAACCB$, $ABABBCBC$, $ABABCBABCA$, $ABABCBCA$, $ABACAACABACAAC$, $ABACABAABACABA$, $ABACABBACABA$, $ABBCAABBCAB$, $ABCAACBA$, $ABCABCBA$, $ABCBACAB$.

Enfin, $ABBCABBC$ et $ABACABAC$ sont 2-DOL-inévitables car simultanément 2-inévitables avec AA et contenus dans les grands carrés, et $ABACBAB$ et $ABCBABC$ sont 3-DOL-inévitables pour la même raison (voir 3.5).

A.5 Résultats d'évitabilité

Nous avons trouvé un grand nombre de motifs évités par le mot de Thue-Morse et une vingtaine d'autres DOL-systèmes binaires simples (voir le dictionnaire pour plus de détails). De même, une douzaine de morphismes de $\{a, b, c\}$

dans $\{a, b\}$, tous appliqués au DOL-langage sans carré classique ($f(a) = abc$, $f(b) = ac$, $f(c) = b$) suffisent à éviter 55 motifs.

Tous ces résultats ont été vérifiés au moyen de l'algorithme présenté au chapitre 6.

A.6 Dictionnaire des motifs

Pour finir, voici un « dictionnaire » contenant tous les motifs ternaires, sauf ceux qui sont 2-DOL-évitables car ils contiennent un motif 2-DOL-évitable figurant dans la liste : le théorème 4.1 assure que cette liste est finie. Pour l'utiliser, il faut tout d'abord mettre le motif étudié p sous *forme normale*, c'est-à-dire permuter les variables de sorte que $p \in A^+B\Delta^*$. On recherche ensuite p dans la liste : s'il y figure, on y lit ce qui est connu sur p. Sinon, p est 2-DOL-évitable; pour trouver le DOL-langage qui évite p, on cherche le plus grand préfixe q de p qui figure dans la liste. Si celui-ci est 2-DOL-évitable, on a trouvé un DOL-langage qui évite p. Sinon, soit p' l'image miroir du préfixe de p de longueur $|q| + 1$: la forme normale de p', ou un de ses préfixes 2-DOL-évitable, devrait être dans la liste.

On a indiqué, pour chaque propriété, comment celle-ci pouvait être démontrée : soit en se ramenant à un autre motif (moyennant parfois le passage par des formules équivalentes : ainsi « $AABCABB$ divise $AABB$ » doit être compris comme « $AABCABB$ est équivalent à $AAB.ABB$ qui divise $AABB$ »), soit par *backtracking* (la longueur maximale des mots qui évitent le motif est alors indiquée, si elle est connue), soit en construisant un L-système (Σ, f, a) ou $(\Sigma, f, a, \Sigma', g)$, auquel cas les morphismes f et éventuellement g sont indiqués sous la forme $f = (abc, ac, b)$ si $f(a) = abc$, $f(b) = ac$ et $f(c) = b$, et les alphabets sont sous-entendus. Pour les motifs k-DOL-inévitables, voir A.4.