Motifs inévitables

Anisse Ismaïli

Math-Info

January 14, 2009

Premières définitions

Définition

Un *motif* est un mot fini sur un alphabet de variables $E = \{\alpha, \beta, \gamma, ...\}$.
Par exemple, $p = \alpha\alpha\beta\beta\alpha$.

Définition

On dit qu'*on rencontre* un motif p dans un mot w sur l'alphabet A, si et seulement si il existe un morphisme $h : E^* \to A^*$ qui substitue chaque variable par un mot non vide de A^*, et tel que $h(p)$ est un facteur de w.

Premières définitions

Exemple

On rencontre le motif $p = \alpha\alpha\beta\beta\alpha$ dans le mot $w = 1011011000111$, par le morphisme $h : \alpha \longmapsto 011, \beta \longmapsto 0$. En effet,

$$w = 1\overbrace{\overbrace{011}^{h(\alpha)}\ \overbrace{011}^{h(\alpha)}\ \overbrace{0}^{h(\beta)}\ \overbrace{0}^{h(\beta)}\ \overbrace{011}^{h(\alpha)}}^{h(\alpha\alpha\beta\beta\alpha)}1$$

Intérêt de l'inévitabilité (1/2)

Nous allons nous intéresser à l'étude de l'inévitabilité des motifs. C'est un sujet passionnant, car étant donné un motif inévitable, on le retrouve absolument partout :

Exemple

Un certain musicien ne connait qu'un nombre fini de notes de musique. On rencontre ce motif inévitable dans tout morceau de musique dépassant une certaine longueur fixe.

Exemple

En Karaté-Do, Il n'y a qu'un nombre fini de mouvements, qu'ils soient d'attaque ou de défense, de jambe ou de bras, gauche ou droit, ... Etant donné un motif inévitable, après avoir fait un certain nombre de mouvement dans un combat, on y rencontre forcément le motif.

Définition de l'inévitabilité

Propriété

Les propriétés suivantes sont équivalentes :

1. Le motif p est inévitable
2. Il existe un rang $n \in \mathbf{N}$ tel que tout mot w de longueur n ou plus contient une occurence de p
3. Tout language infini $L \in P(A^*)$ contient une occurence de p.
4. Tout mot infini $u \in A^\omega$ contient une occurence de p.

Intéret de l'inévitabilité (2/2)

On peut se poser la question de l'utilité de ce concept abstrait. C'est utilisable pour compresser des données : Pas besoin de stocker tout le fichier : On ne garde que le motif, et le mot non-vide associé à chaque variable, d'où un gain de taille sur le fichier.

Exemple

On se donne un motif p qui est inévitable. Considérant un mot $w = uh(p)v$ (ou un fichier w...) où on recontre le motif p par le morphisme h, la connaissance de u, v, p et h, permet un représentation plus condensée de w.

Exemple

Grossièrement, la connaissance des motifs inévitables peut indiquer à notre Karateka une idée de quels enchainements il doit travailler.

Motifs k-inévitables et motifs inévitables

Dans le cadre de cette étude passionnante, nous allons distinguer différentes notions : la k-évitabilité, la k-inévitabilité, l'évitabilité et l'inévitabilité.

Motifs k-évitables et motifs k-inévitables

Définition

Un motif est *k-évitable* si et seulement si il est évitable sur n'importe quel alphabet à k lettres. Un motif qui n'est pas k-évitable est *k-inévitable*.

Exemple

Le motif $p = \alpha\alpha = \alpha^2$ est 2-inévitable, car tout mot de longueur au moins 4, sur un alphabet binaire, contient au moins un carré. Mais p est 3-évitable. En effet, $u = abcacbabcbac...$ le point fixe du morphisme ternaire $\mu : a \longmapsto abc, b \longmapsto ac, c \longmapsto b$ évite le motif α^2, qui est donc 3-évitable.

Motifs évitables et motifs inévitables

Définition

Un motif qui est évitable sur un certain A sera simplement dit *évitable*. Un motif qui est inévitable pour tout A sera simplement dit *inévitable*.

Exemple

Grace à l'algorithme de Zimin, on verra plus tard que $p = \alpha\beta\alpha\gamma\xi\beta\alpha\delta\alpha\beta\xi\gamma\xi\beta\xi$ est inévitable.

Hiérarchisation

Propriété

On a la hiérarchisation suivante :

$$\{\text{2-évitable}\} \subset \{\text{3-évitable}\} \subset \ldots \subset \{\text{k-évitable}\} \subset \ldots \subset \{\text{évitable}\}$$
$$\{\text{2-inévitable}\} \supset \ldots \supset \{\text{k-inévitable}\} \supset \ldots \supset \{\text{inévitable}\}$$

Indice d'évitabilité

Définition

L'*indice d'évitabilité* $\mu(p)$ d'un motif $p \in E^*$ est le plus petit entier k tel que p soit k-évitable, ou ∞ si p est inévitable.

Exemple

α^2 est 2-inévitable mais 3-évitable. Donc $\mu(\alpha^2) = 3$.

Propriété

Clairement, on a : $2 \leq \mu(p) \leq \infty$ et d'une autre part, si $p|q$ alors $\mu(p) \geq \mu(q)$.

Exemples de motifs : Puissances

La classe des puissances α^n d'une variable unique est la classe de motifs la plus simple. $\alpha^0 = \epsilon$ et $\alpha^1 = \alpha$ sont trivialement inévitables, puisqu'elles sont rencontrées dans n'importe quel mot non-vide. α^2 est 2-inévitable mais 3-évitable. Et pour $n \geq 3$, α^n est 2-évitable.

Exemples de motifs : Motifs de Zimin (1/3)

Propriété

Le motif $\alpha\beta\alpha$ est inévitable. Plus précisemment, si $|A| = k$, tout mot de longueur au moins 2k+1 contient une occurence de $\alpha\beta\alpha$, et cette borne est serrée.

Preuve. Soit $w \in A^*$ un mot de longueur au moins $2k+1$. Alors une des lettres de A (comme nous sommes à permutation de A près, nous pouvons dire a), apparait au moins 3 fois dans w. On écrit $w = w_0 a w_1 a w_2 a w_3$. Soit alors $h : (h(\alpha) = a, h(\beta) = w_1 a w_2)$ est un morphisme non-effaçant et $h(\alpha\beta\alpha)$ est un facteur de w. Si $A = \{a_1, a_2, ..., a_k\}$, alors $w = a_1 a_1 a_2 a_2 ... a_k a_k$ est un mot de longueur $2k$ qui évite $\alpha\beta\alpha$.

Exemples de motifs : Motifs de Zimin (2/3)

Propriété

Soit p un motif inévitable sur A, et ζ une variable qui n'a pas d'occurences dans p. Alors le motif $p\zeta p$ est inévitable sur A.

Preuve. Soit $k = |A|$. Comme p est inévitable sur A, il existe un entier l tel que tout mot $w \in A^l$ contient p. A^l est un ensemble fini de k^l mots. Soit $N = k^l(l+1) + l$, et soit $w \in A^N$. Le mot w peut être vu comme la concaténation de $k^l + 1$ mots de longueur l, séparé par de simples lettres. Parmis ces $k^l + 1$ facteurs de longueur l, au moins deux sont égaux (en vertue du théorème des $k^l + 1$ chaussettes dans k^l tiroirs à chaussettes), disons v. On écrit alors $w = w_0 v w_1 v w_2$ avec $|v| = l$ et $|w_1| \geq 1$. Comme v est de longueur l, il rencontre le motif p. Ainsi, il existe un morphisme non-effaçant $h : (alphabet(p))^* \to A^*$ tel que $v = v_0 h(p) v_1$. D'où $w = w_0 v_0 h(p) v_1 w_1 v_0 h(p) v_1 w_2$. En posant $h(\zeta) = v_1 w_1 v_0$, on trouve que $h(p\zeta p)$ est un facteur de w.

Exemples de motifs : Motifs de Zimin (3/3)

Définition

Si on applique la proposition précédente en partant du mot vide, on peut construire une famille infinie de motifs inévitables. Soient $(\alpha_n)_{n\in\mathbb{N}}$ différentes variables de E. Soit $Z_0 = \epsilon$ et pour tout $n \in \mathbb{N}$, $Z_{n+1} = Z_n\alpha_n Z_n$. Les motifs Z_n sont appelés *mots de Zimin*.

Propriété

Les motifs de Zimin Z_n sont tous inévitables.

Preuve. Soit A un alphabet fini. Nous avons vu que $Z_0 = \epsilon$, $Z_1 = \alpha_0$ et $Z_2 = \alpha_0\alpha_1\alpha_0$ sont inévitables sur A. Si Z_n est inévitable sur A, d'après la proposition précédente, $Z_{n+1} = Z_n\alpha_n Z_n$ est aussi inévitable. Comme toutes les motifs de zimin Z_n sont inévitables sur tout A, ils sont inévitables.

Théorème principal

Pour montrer que l'inévitabilité est décidable, nous allons montrer qu'elle est équivalente à une certaine propriété de *réductibilité*, définie ci-dessous, et qui peut elle-même être vérifiée par un algorithme récursif. Définissons d'abord le procédé de réduction.

Théorème

Un motif est inévitable si et seulement si il est réductible.

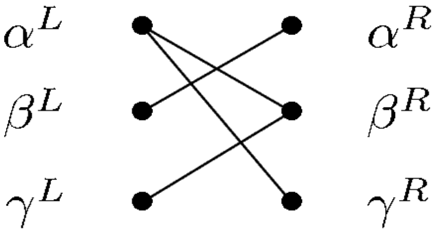
Graphe d'adjacence

Définition

Soit $p \in E^*$ un motif. Le *graphe d'adjacence* de p est le graphe biparti $AG(p)$ avec deux copies de E comme sommets, E^L (pour E left) et E^R (pour E right), et une arête entre ξ^L et η^R si et seulement si $\xi\eta$ est un facteur de p.

Exemple

Le graphe d'adjacence de $\alpha\beta\alpha\gamma\beta\alpha$.



Parties Libres

Définition

Un sous-ensemble F de *alphabet*$(p) \subset E$ est appelé *partie libre* si il n'existe pas de chemin dans $AG(p)$ reliant un sommet gauche ξ^L à un sommet droit η^R, avec ξ et η dans F. Pour trouver toutes les parties libres, il faut determiner les composantes connexes de $AG(p)$.

Exemple

Dans l'exemple précédent, $AG(\alpha\beta\alpha\gamma\beta\alpha)$ a deux composantes connexes ($\{\alpha^R, \beta^L\}$ et $\{\alpha^L, \beta^R, \gamma^L, \gamma^R\}$), et deux parties libres, $\{\alpha\}$ et $\{\beta\}$.
Aucune partie libre ne contient γ car il existe un chemin de γ^L à γ^R.
Aucune partie libre ne contient α ET β car il existe un chemin de β^L à α^R.

Réduction

Définition

- Étant donné un motif p et une partie libre F de p, on dit que p se *réduit en une étape à q par la suppression de F*, si q est le motif obtenu en supprimant toute les occurence des lettres de F dans p. On le notera $p \xrightarrow{F} q$.
- On dit que p *se réduit à q* si il existe une suite de réductions à une étape, allant de p à q. On le notera $p \xrightarrow{*} q$.
- Enfin, un motif p est *réductible* si il se réduit au motif vide, $p \xrightarrow{*} \epsilon$. Sinon p est *irréductible*.

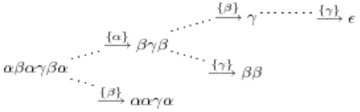
Remarque

Pour montrer qu'un motif est irréductible, il est nécessaire d'explorer recursivement toutes les possibilités de réductions à une étape, pour s'assurer qu'aucune d'elles ne mène au motif vide.

Exemple de réduction (1/2)

Exemple

Arbre des réductions de $p = \alpha\beta\alpha\gamma\beta\alpha$:



Ainsi, p se réduit à ϵ par la suite de réductions $\alpha\beta\alpha\gamma\beta\alpha \xrightarrow{\{\alpha\}} \beta\gamma\beta \xrightarrow{\{\beta\}} \gamma \xrightarrow{\{\gamma\}} \epsilon$. On montrera par la suite que cela implique qu'il est inévitable.

Exemple de réduction (2a/2)

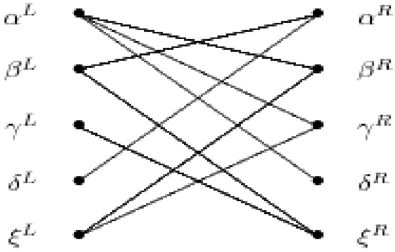
Remarque

Il est parfois nécessaire de supprimer des parties libres de plus d'un élément, comme par exemple dans l'exemple suivant.

Exemple de réduction (2b/2)

Exemple

Graphe d'adjacence du motif $p = \alpha\beta\alpha\gamma\xi\beta\alpha\delta\alpha\beta\xi\gamma\xi\beta\xi$, dont on déduit les composantes connexes $\{\alpha^R, \beta^L, \gamma^L, \delta^L, \xi^R\}$ et $\{\alpha^L, \beta^R, \gamma^R, \delta^R, \xi^L\}$, et les parties libres $\{\alpha\}$, $\{\beta\}$, $\{\gamma\}$, $\{\delta\}$, $\{\xi\}$, $\{\alpha, \xi\}$, $\{\beta, \gamma\}$, $\{\beta, \delta\}$, $\{\gamma, \delta\}$, $\{\beta, \gamma, \delta\}$:



Exemple de réduction (2c/2)

Illustrant la remarque précédente, supprimer un singleton conduit à des motifs irréductibles, tandis que commencer par supprimer la partie libre $\{\alpha, \xi\}$ permet la reduction :

$$\alpha\beta\alpha\gamma\xi\beta\alpha\delta\alpha\beta\xi\gamma\xi\beta\xi \xrightarrow{\{\alpha,\xi\}} \beta\gamma\beta\delta\beta\gamma\beta \xrightarrow{\{\beta\}} \gamma\delta\gamma \xrightarrow{\{\gamma\}} \delta \xrightarrow{\{\delta\}} \epsilon$$

On peut aussi remarquer que $\beta\gamma\beta\delta\beta\gamma\beta$ est un motif de Zimin, donc inévitable, *donc* réductible.

l'algorithme récursif de Zimin

ENTRÉES : $p \in E^*$
SORTIES : rend vrai ssi p est réductible
si $p = \epsilon$ **alors** Retourner **VRAI**
$AG(p) \leftarrow \emptyset$
$Sommets(AG(p)) \leftarrow alphabet(p) \times \{L, R\}$
pour tout $\xi\eta \in facteurs(p)$ **faire**
    $Arretes(AG(p)) \leftarrow Arretes(AG(p)) \cup \{\{\xi_L, \eta_R\}\}$
fin pour
$\{F_i\}_{i \in I} \leftarrow PartiesLibres(AG(p))$
pour tout $F_i \in \{F_i\}_{i \in I}$ **faire**
    $q \xleftarrow{F_i} p$
    si $AlgorithmeZimin(q)$ **alors** Retourner **VRAI**
fin pour
Retourner **FAUX**

Introduction de la classification binaire

- Contrairement au cas de l'algorithme de Zimin, où la taille de l'alphabet n'avait pas d'importance, il n'existe pas d'algorithme connu pour déterminer l'indice dévitabilité d'un motif. Par exemple, à l'heure actuelle, on ne sait pas si $\mu(\alpha\alpha\beta\beta\gamma\gamma)$ vaut 2 ou 3.
- Le cas unaire a déjà été traité plus haut. On a : $\mu(\epsilon) = \mu(\alpha) = \infty$, $\mu(\alpha\alpha) = 3$ et pour $k \geq 3$, $\mu(\alpha^k) = 2$.
- Restreignons nous donc au cas des motifs binaires. i.e.:$E = \{\alpha, \beta\}$

Motifs binaires inévitables

Le fait que $\alpha\alpha$ soit 3-évitable et 2-inévitable nous donne déjà beaucoup d'informations sur les motifs binaires. Seul un nombre fini de motifs sont 3-inévitables. En effet, un motif divisible par $\alpha\alpha$ doit être 3-évitable, et comme $\alpha\alpha$ est 2-inévitable, il n'y a qu'un nombre fini de motifs qui ne sont pas divisibles par $\alpha\alpha$, à savoir ϵ, α, β, $\alpha\beta$, $\beta\alpha$, $\alpha\beta\alpha$, et $\beta\alpha\beta$. En fait, ils sont tous inévitables, ce qui implique que l'indice dévitabilité d'un motif binaire ne peut être que 2, 3, où ∞.

Motifs binaires d'indice 3 (1/2)

Théorème

Les motifs binaires $\alpha\alpha$, $\alpha\alpha\beta$, $\alpha\alpha\beta\alpha$, $\alpha\beta\beta\alpha$, $\alpha\alpha\beta\beta$, $\alpha\beta\alpha\beta$, $\alpha\alpha\beta\alpha\alpha$ et $\alpha\alpha\beta\alpha\beta$ sont d'indice d'évitabilité 3.

Preuve. Comme tout ces motifs sont divisibles par $\alpha\alpha$, ils sont 3-évitables. Un simple algorithme est suffisant pour montrer qu'ils sont 2-inévitables. Les résultats des algorithmes sont résumés dans le tableau suivant, où w est un exemple de mot binaire évitant p et de longueur maximale, et N est le nombre total de mots évitant p, en incluant le mot vide et les mots unaires.

Motifs binaires d'indice 3 (2/2)

p	w	$\lvert w \rvert$	N
$\alpha\alpha$	aba	3	7
$\alpha\alpha\beta$	$abab$	4	13
$\alpha\alpha\beta\alpha$	$ababab aaa$	9	91
$\alpha\beta\beta\alpha$	$aabbbaaabb$	10	93
$\alpha\alpha\beta\beta$	$abaaabaaaba$	11	147
$\alpha\beta\alpha\beta$	$abaabbaaabbbaabbab$	18	477
$\alpha\alpha\beta\alpha\alpha$	$abaaaabbbabababab$	18	1699
$\alpha\alpha\beta\alpha\beta$	$abababab aabbaaabbaaaabbbbaaabbbaabbab$	38	26241

Motifs binaires d'indice 2

Théorème

Les motifs binaires $\alpha\alpha\alpha$, $\alpha\beta\alpha\beta\alpha$, $\alpha\beta\alpha\beta\beta\alpha$, $\alpha\alpha\beta\alpha\beta\beta$, $\alpha\beta\alpha\alpha\beta$, et $\alpha\alpha\beta\beta\alpha$ ont un indice d'évitabilité d'index 2.

- *Preuve.* Les motifs $\alpha\alpha\alpha$ et $\alpha\beta\alpha$ sont évités par le mot infini de Thue-Morse, et par conséquent 2-évitables.
- Les deux motifs $\alpha\beta\alpha\beta\beta\alpha$ et $\alpha\alpha\beta\alpha\beta\beta$ sont évités par le mot infini $u = \nu^\omega(a)$, où ν est le morphisme uniforme $a \mapsto aab$, $b \mapsto bba$.
- Le motif $\alpha\beta\alpha\alpha\beta$ est évité par le mot infini $v = \psi(\mu^\omega(a))$, où μ est le morphisme ternaire défini par $a \mapsto abc$, $b \mapsto ac$ et $c \mapsto b$, et ψ par $a \mapsto aaa$, $b \mapsto bbb$ et $c \mapsto ababab$.
- Le motif $\alpha\alpha\beta\beta\alpha$ est évité par le mot infini $v = \chi(\mu^\omega(a))$, où μ est définie comme précédemment et χ par $a \mapsto aa$, $b \mapsto aba$ et $c \mapsto abbb$.

Arbre de la Classification