

UNIVERSITÉ PARIS 6  
THÈSE DE DOCTORAT  
Spécialité: Informatique

présentée par

Julien Cassaigne

pour obtenir le grade de Docteur de l'Université Paris 6

Sujet de la thèse :

*Motifs évitables et régularités dans les mots*

soutenue le 10 juin 1994 devant le jury composé de MM.

Jean-Paul Allouche	(rapporteur)
Jean Berstel	(directeur)
Maxime Crochemore	(rapporteur)
Pavel Goralčik	
Michel Mendès-France	
Jean-Éric Pin	(président)

## Résumé

Nous étudions dans cette thèse les motifs évitables sur un alphabet fixé, ainsi que d'autres régularités évitables ou non. En particulier, nous présentons une classification complète des motifs binaires et un début de classification des motifs ternaires; nous prouvons par ailleurs que l'ensemble des motifs inévitables sur un alphabet donné et ayant un nombre de variables donné est toujours fini. Nous introduisons les notions de DOL-évitabilité et de HDOL-évitabilité, qui traduisent le lien entre l'évitabilité des motifs et les L-systèmes, et nous montrons que sous certaines conditions on peut décider si un HDOL-langage évite un motif. Nous abordons également deux problèmes plus combinatoires, celui du dénombrement des mots évitant un motif, dans le cas particulier des chevauchements, et celui de la construction de suites de complexité donnée.

**Mots-clés :** Théorie des langages, Combinatoire, Mot, Motif évitable, Décidabilité, L-système, Chevauchement, Complexité de suites, Suite régulière, Facteur spécial

## Abstract

### *Avoidable patterns and regularities in words*

In this thesis, we study the avoidability of patterns on a fixed alphabet, and other regularities in words, avoidable or unavoidable. Specifically, we give a complete classification of binary patterns and a partial classification of ternary patterns; moreover, we prove that there are always finitely many unavoidable patterns on a given alphabet with a given number of variables. We introduce DOL-avoidability and HDOL-avoidability, which reflect the link between pattern avoidability and L-systems, and we show that it is possible, with a few restrictions, to decide whether a HDOL-language avoids a pattern. We also consider two more combinatorial problems, namely counting words avoiding a pattern, in the special case of overlaps, and constructing sequences with a given subword complexity.

**Keywords:** Language theory, Combinatorics, Word, Avoidable pattern, Decidability, L-system, Overlap, Subword complexity, Regular sequence, Special factor

# Table des matières

<b>Introduction</b>	<b>5</b>
<b>1 Mots et motifs</b>	<b>9</b>
1.1 Définitions usuelles . . . . .	9
1.2 Motifs . . . . .	10
1.3 Évitaibilité . . . . .	11
1.3.1 Définition . . . . .	11
1.3.2 Indice d'évitaibilité . . . . .	11
1.3.3 Exemples . . . . .	12
1.3.4 $\mathcal{L}$ -évitaibilité . . . . .	12
1.4 Premières propriétés . . . . .	13
1.4.1 Symétries . . . . .	13
1.4.2 Divisibilité . . . . .	13
1.4.3 Langages répétitifs et motifs . . . . .	13
1.5 Motifs avec constantes, formules . . . . .	14
1.5.1 Motifs avec constantes . . . . .	14
1.5.2 Ensembles de motifs . . . . .	14
1.5.3 Formules . . . . .	15
1.5.4 Répétitions . . . . .	20
1.6 Variantes . . . . .	20
1.6.1 Évitaibilité commutative . . . . .	20
1.6.2 $l$ -occurrences . . . . .	21
1.7 Évitaibilité sur un alphabet non fixé . . . . .	22
1.7.1 Quelques propriétés . . . . .	22
1.7.2 L'algorithme de Zimin-Bean-Ehrenfeucht-McNulty . . . . .	23
1.8 Motifs verrouillés . . . . .	24
<b>2 Classification des motifs binaires</b>	<b>29</b>
2.1 Les motifs inévitables . . . . .	29
2.2 Motifs 2-inévitables . . . . .	29
2.2.1 L'algorithme de <i>backtracking</i> . . . . .	29
2.2.2 Résultats . . . . .	30
2.2.3 Accélération du <i>backtracking</i> . . . . .	31
2.3 3-évitaibilité . . . . .	33

2.4	2-évitabilité: les résultats de P. Roth . . . . .	33
2.5	Le cas de $AABBA$ . . . . .	34
2.6	Autres preuves de 2-évitabilité . . . . .	39
2.7	Classification des motifs binaires . . . . .	41
<b>3</b>	<b>DOL-évitabilité, HDOL-évitabilité</b>	<b>43</b>
3.1	DOL-langages et HDOL-langages . . . . .	43
3.1.1	DOL-langages . . . . .	44
3.1.2	HDOL-langages . . . . .	44
3.1.3	Définitions et propriétés usuelles . . . . .	44
3.2	Motifs DOL-évitables et HDOL-évitables . . . . .	45
3.3	Quelques résultats faciles . . . . .	45
3.4	Réduction des DOL et HDOL . . . . .	46
3.4.1	Minimalité de l'alphabet . . . . .	48
3.4.2	Élimination de l'axiome . . . . .	48
3.4.3	Morphismes élémentaires . . . . .	49
3.4.4	Itération du morphisme . . . . .	50
3.4.5	Morphismes non-effaçants . . . . .	52
3.4.6	Le cas des L-systèmes <i>pushy</i> . . . . .	53
3.4.7	Simplification des images des lettres de rang 0 . . . . .	55
3.4.8	Morphismes expansifs . . . . .	55
3.4.9	Récapitulation . . . . .	57
3.5	Preuves de DOL-inévitabilité . . . . .	57
3.6	Nouvelle classification des motifs binaires . . . . .	60
<b>4</b>	<b>Finitude de l'ensemble des motifs <math>n</math>-aires <math>k</math>-inévitables (<math>k \geq 2</math>)</b>	<b>63</b>
4.1	Introduction: $\ell_{nk}$ . . . . .	63
4.2	Les grands motifs $n$ -aires sont 2-DOL-évitables . . . . .	64
4.3	Une meilleure borne avec des HDOL-systèmes . . . . .	67
4.4	Valeurs connues de $\ell_{nk}$ . . . . .	69
<b>5</b>	<b>Dénombrément des mots évitant un motif:</b>	
	<b>le cas des chevauchements</b>	<b>71</b>
5.1	Introduction . . . . .	71
5.2	Résultats antérieurs . . . . .	72
5.3	Une bijection entre le langage des mots binaires sans chevauchement et un langage rationnel . . . . .	75
5.3.1	Une décomposition des mots binaires sans chevauchement . . . . .	75
5.3.2	Les langages rationnels qui représentent $U$ et $V$ . . . . .	77
5.4	Des matrices pour calculer $u_n$ . . . . .	80
5.5	Étude asymptotique de la suite $(u_n)$ . . . . .	83
5.5.1	Simplification des relations de récurrence . . . . .	83
5.5.2	Les exposants $\alpha$ et $\beta$ . . . . .	84
5.5.3	Sommes partielles . . . . .	86

5.6	Avec d'autres motifs . . . . .	87
<b>6</b>	<b>Un algorithme pour décider si un HDOL évite un motif</b>	<b>89</b>
6.1	Morphismes circulaires . . . . .	89
6.1.1	Morphismes circulaires sur $\Sigma^*$ . . . . .	89
6.1.2	Morphismes circulaires sur $U$ . . . . .	91
6.1.3	L-systèmes circulaires . . . . .	92
6.2	L'image inverse d'un motif par un morphisme circulaire . . . . .	93
6.3	L'algorithme . . . . .	96
6.4	Cas des morphismes non expansifs . . . . .	98
6.5	Conclusion . . . . .	100
6.6	Exemple . . . . .	101
<b>7</b>	<b>Une autre application des morphismes circulaires : calculs de complexités de suites</b>	<b>105</b>
7.1	Introduction . . . . .	105
7.2	Suites ou langages? . . . . .	106
7.3	Facteurs spéciaux et bispéciaux . . . . .	106
7.3.1	Facteurs spéciaux . . . . .	106
7.3.2	Facteurs bispéciaux . . . . .	107
7.3.3	Le cas des alphabets plus grands . . . . .	110
7.4	Action d'un morphisme circulaire . . . . .	112
7.5	Application: suites de complexité affine . . . . .	113
7.5.1	Suites de complexité ultimement affine . . . . .	113
7.5.2	Complexité $\alpha n + \beta$ pour $n \geq 1$ . . . . .	115
7.6	Quelques exemples de calculs de complexité de suites définies par morphisme . . . . .	116
7.6.1	Suite de Thue-Morse . . . . .	116
7.6.2	Un exemple non primitif . . . . .	117
	<b>Conclusion</b>	<b>119</b>
<b>A</b>	<b>Classification partielle des motifs ternaires</b>	<b>121</b>
A.1	Les résultats de Silke Nilgens . . . . .	121
A.2	Les 5 motifs restants . . . . .	122
A.3	Résultats de 2-inévitabilité . . . . .	123
A.4	Résultats de DOL-inévitabilité . . . . .	124
A.5	Résultats d'évitabilité . . . . .	124
A.6	Dictionnaire des motifs . . . . .	125
	<b>Bibliographie</b>	<b>161</b>
	<b>Notations</b>	<b>167</b>
	<b>Index</b>	<b>169</b>