Les motifs évitables

Année 2015-2016

Auteurs : CAPITAINE Théo, NICOLLE Clément et RICHET Nicolas, élèves de Terminale S.
Établissement : Lycée Condorcet, Saint-Quentin (02).
Encadrés par : Fabien AOUSTIN.
Chercheur : Fabien DURAND, Laboratoire Amiénois de Mathématique Fondamentale et Appliquée, LAMFA, UMR CNRS 7352, Université de Picardie Jules Verne.

Présentation du sujet

Le but de ce sujet est de créer un mot infini à l'aide d'un alphabet défini tout en évitant la répétition d'un certain motif.

Annonce des conjectures et résultats obtenus

Résultats obtenus :

- Il est impossible de créer un mot infini à l'aide d'un alphabet de deux caractères évitant le motif carré.
- Il est possible de créer un mot infini à l'aide d'un alphabet de deux caractères évitant le motif cube.
- Il est possible de créer un mot infini à l'aide d'un alphabet de trois caractères évitant le motif carré.

1. Définition des notions :

Alphabet : Un alphabet est un ensemble de caractères utilisés afin de créer un mot.
Motif : Un motif est une répétition identique et successive d'un ensemble de caractères de cet alphabet. Ici on nommera deux motifs différents : **le carré** (répétition de deux caractères ou de deux ensembles de caractères identiques et successifs) et **le cube** (répétition de trois caractères ou de trois ensembles de caractères identiques et successifs).

Exemples : Sur l'alphabet {*a* ; *b*} on peut considérer les mots :

- *aba ;*
- *aab* (qui contient un motif carré) ;
- *ababa* (qui contient un motif carré) ;
- *bbababaabab* (qui contient un motif cube) ;

2. Un mot infini à deux caractères sans carré ?

Premièrement, on s'est demandé s'il était possible de créer un mot infini à l'aide d'un alphabet de deux caractères évitant le motif carré.
Pour fixer les choses, utilisons l'alphabet {*a*, *b*}.
En y réfléchissant un peu, on se rend assez vite compte que cela va être compliqué.
Démonstration :
On peut commencer notre mot soit par *a*, soit par *b*.
Ici on fait le choix de commencer par *a* mais cela n'a aucune importance :

a

Afin d'éviter le motif *aa* qui est un carré, nous sommes dans l'obligation de devoir ajouter un *b* :

ab

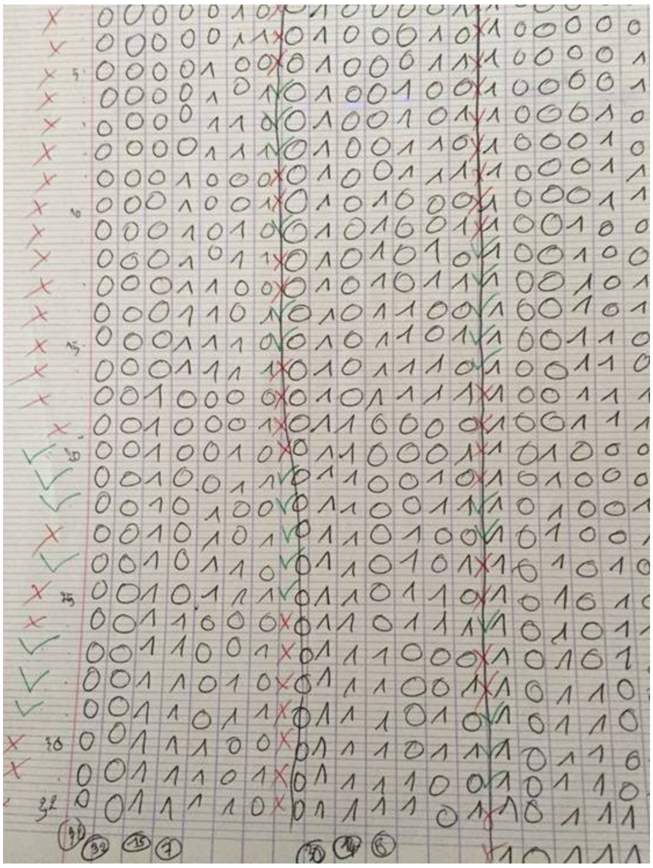
De la même façon, nous nous retrouvons dans l'obligation d'ajouter un *a* afin d'éviter le motif *bb* :

aba

À cette étape, nous sommes coincés car si on ajoute un *a* nous avons le motif *aa*, et si nous ajoutons un *b* nous avons le motif *abab*.
Il est donc impossible de créer un mot infini à l'aide d'un alphabet de deux caractères évitant le motif carré.

3. Quelques statistiques :

Pour obtenir quelques statistiques nous avons d'abord énuméré à la main les mots de longueur 3, 4, 5, 6, etc. sur l'alphabet {0, 1} puis compté le nombre de mots avec ou sans cube. Voici une de nos feuilles de brouillon :



Pour aller plus loin, nous avons créé un algorithme (en C++) permettant de générer des nombres de *n* caractères et de comparer chaque caractère du nombre, nous indiquant ainsi la présence d'un motif cube ou non.
Nous avons fait fonctionner cet algorithme jusqu'à la création de mots de 20 caractères et il nous a généré les 2^{20} possibilités correspondant à tous les mots possibles avec 2 caractères différents.

Nous avons réunis tous ces résultats dans le tableau ci-dessous :

Nombre de caractères par mot	1	2	3	4	5	6	7	8	9
Nombre de mots	2	4	8	16	32	64	128	256	512
Nombre de mots sans cube	2	4	6	10	16	24	36	56	80
Nombre de mots avec cube(s)	0	0	2	6	16	40	92	200	432

10	11	12	13	14	15	16	17	18	19	20
1024	2048	4096	8192	16384	32768	65536	131072	262144	524288	1048576
118	174	254	378	534	802	1168	1716	2502	3650	5324
906	1874	3842	7814	15850	31966	64368	129356	259642	520638	1043252

Nous avons donc remarqué que le nombre de mots avec cube augmente mais aussi le nombre de mots sans cube, nous avons ainsi conjecturé que nous pouvions ainsi sûrement créer un mot aussi grand que l'on veut à 2 caractères différents en évitant les cubes.
Si on note u_n le nombre de mots sans cubes à n lettres sur $\{0, 1\}$, on a remarqué quelques relations intéressantes :

$u_3 = u_2 + u_1$ $\quad$ $u_7 = u_6 + u_5 - 4$

$u_4 = u_3 + u_2$ $\quad$ $u_8 = u_7 + u_6 - 4$

$u_5 = u_4 + u_3$ $\quad$ $u_9 = u_8 + u_7 - 12$

$u_6 = u_5 + u_4 - 2$ $\quad$ $u_{10} = u_9 + u_8 - 18$

Hélas, cela ne nous a rien donné de plus.

4. Construction du mot de Morse :

Le mot de Morse est un mot construit sur l'alphabet $\{0, 1\}$ grâce à un morphisme σ.
Le morphisme σ est défini par $\sigma(0) = 01$ et $\sigma(1) = 10$.
On commence par le mot « 0 » et on applique le morphisme plusieurs fois.
On obtient alors la suite de mots suivante :

0
01
0110
01101001
0110100110010110
01101001100101101001011001101001
0110100110010110100101100110100110010110011010010110100110010110
etc.

On peut remarquer qu'un mot de Morse à l'étape n est constitué du mot de Morse à l'étape $n-1$ collé à son conjugué.
Par exemple, le conjugué de 0110 est 1001 et le mot de Morse suivant est 01101001.
Nous avons observé qu'il n'y avait pas de cube dans les mots ainsi obtenus.
Nous avons aussi observé qu'on ne trouve jamais de motif de la forme $uu\varepsilon$ où u est un mot et où ε est un caractère.
Nous allons démontrer cela dans le paragraphe suivant ce qui montrera qu'il n'y a pas de cube dans les mots de Morse.

5. Mot infini à deux caractères sans cube ?

On note $u^{2+\varepsilon}$ un motif du type $uu\varepsilon$ où u est un mot et où ε est une lettre.

À partir de cette étape on cherche à démontrer que le motif $u^{2+\varepsilon}$ ne peut pas apparaitre dans un mot de Morse. On choisit de passer par un raisonnement par l'absurde. On adopte ainsi une décomposition pour un mot de Morse dans laquelle on fait apparaitre le motif :

$$w \; \varepsilon \; v \; \varepsilon \; v \; \varepsilon \; x$$

Dans cette écriture, w, v et x représentent des mots quelconques et ε représente un caractère. Ainsi ici on voit que le motif $u^{2+\varepsilon}$ « $\varepsilon v \varepsilon v \varepsilon$ » apparait.

On sait qu'un mot de Morse possède un nombre de caractères pair, on en déduit que le mot $wvvx$ possède un nombre de caractères impair.

On peut ainsi regrouper les mots « v » ensemble, la parité de leur nombre de caractères seront donc sans importance car, étant additionnés ensemble, leur nombre de caractères sera toujours pair.

Le mot wx possède donc un nombre de caractères impair, ce qui signifie que les mots w et x ont un nombre de caractères de parité différente.

Cela va nous aider dans la simplification de ce mot puisque cela signifie que si on cherche à commencer par un mot pair ou impair, on a juste à lire le mot de Morse dans le sens inverse, celui-ci étant composé d'un mot et de son conjugué.

On peut donc supposer que w est de longueur paire et on en arrive au mot : $\varepsilon v \varepsilon v \varepsilon$.

On remarque qu'il y a deux cas à traiter, le cas où v est de longueur paire et le cas où v est de longueur impaire.

On commence par la démonstration pour le cas où v est de longueur paire.

On utilise une méthode par substitution. On donne un exemple de cette méthode mais celle-ci fonctionne dans le cas général :

On suppose que v est composé de 6 caractères, il peut donc s'écrire sous la forme :

$$\underbrace{\varepsilon \; n_1} \; \underbrace{n_2 \; n_3} \; \underbrace{n_4 \; n_5} \; \underbrace{n_6 \; \varepsilon} \; \underbrace{n_1 \; n_2} \; \underbrace{n_3 \; n_4} \; \underbrace{n_5 \; n_6} \; \varepsilon$$

À cette étape, on regroupe les caractères par deux, ce qui nous permet de créer des égalités entre certains groupes. Par construction du mot de Morse, et comme on a supposé w de longueur paire, chaque groupe est égal à « 01 » ou « 10 ».

Les groupes εn_1 et $n_6 \varepsilon$ sont donc des groupes conjugués, on en déduit que $n_1 = n_6$. En injectant ceci dans les groupes de droite, on obtient de la même façon que $n_2 = n_5$. On utilise le même procédé en injectant à chaque fois l'égalité de l'autre côté afin de retrouver une nouvelle égalité. À la fin, on obtient un groupe, ici $n_3 n_4$, où les deux caractères sont les mêmes ce qui est impossible.

On en déduit que v de peut pas être de longueur paire.

Passons maintenant au cas où v est de longueur impaire.

Si le mot de Morse à l'étape n est de la forme $w \; \varepsilon \; v \; \varepsilon \; v \; \varepsilon \; x$ et que v est de longueur impaire, alors le mot de Morse à l'étape $n-1$ est de la forme $w' \; \varepsilon \; v' \; \varepsilon \; v' \; \varepsilon \; x'$.

En suivant le même raisonnement que précédemment, on peut supposer que w' est de longueur paire. On a déjà vu dans le cas précédent que v' ne peut pas être de longueur paire. Ainsi, v' est de longueur impaire.

On peut répéter ce processus jusqu'à remonter à un mot de Morse contenant un motif de la forme $\varepsilon\bar{\varepsilon}\varepsilon\bar{\varepsilon}\varepsilon$.

En remontant encore d'une étape, on arrive à un mot de Morse contenant le motif $\varepsilon\varepsilon\varepsilon$, ce qui n'est pas possible.

On constate que v ne peut être ni de longueur paire, ni de longueur impaire. Il n'y a donc jamais de motif $u^{2+\varepsilon}$ dans un mot de Morse et donc aucun cube dans le mot de Morse infini.

6. Et les carrés sur trois caractères ?

Nous avons aussi cherché à la main si on pouvait construire un mot sur $\{0, 1, 2\}$ sans carré et aussi grand que l'on veut.
En notant u_n le nombre de mots de n lettres sans carrés sur $\{0, 1, 2\}$ on a obtenu :

n	2	3	4	5	6	7	8	9	10
u_n	6	12	18	30	42	60	78	108	150

Nous avons remarqué quelques relations :

$$u_4 = u_2 + u_3$$
$$u_5 = u_3 + u_4$$
$$u_6 = u_3 + u_5$$
$$u_7 = u_4 + u_6$$
$$u_8 = u_4 + u_7$$
$$u_9 = u_5 + u_8$$
$$u_{10} = u_6 + u_9$$

Nous avons conjecturé que u_n était toujours la somme de deux termes précédents mais nous n'avons pas réussi à le démontrer.

En partant du mot de Morse infini, nous avons trouvé un moyen d'obtenir un mot infini sans carré sur $\{0, 1, 2\}$.
Dans le même esprit que la construction du mot de Morse, l'élaboration de ce mot permettant d'éviter les carrés avec l'alphabet $\{0, 1, 2\}$ se fait grâce à un morphisme σ'.
Le morphisme σ' est défini par $\sigma'(011) = 2$, $\sigma'(01) = 1$ et $\sigma'(0) = 0$.
On lit le mot de Morse infini de gauche à droite.
Si on rencontre le motif « 011 », on le remplace par un « 2 ».
Sinon, si on rencontre le motif « 01 », on le remplace par « 1 ».
Si on n'est dans aucun des deux cas précédents, on tombe forcément sur le motif « 0 » et on le remplace par « 0 ».
Voilà ce que donne cette transformation pour le début du mot de Morse infini

$$\underbrace{011}_{2}\underbrace{01}_{1}\underbrace{0}_{0}\underbrace{011}_{2}\underbrace{0}_{0}\underbrace{01}_{1}\underbrace{011}_{2}\underbrace{01}_{1}\underbrace{0}_{0}\underbrace{01}_{1}\underbrace{011}_{2}\underbrace{0}_{0}\underbrace{011}_{2}\underbrace{01}_{1}\underbrace{0}_{0}\underbrace{011}_{2}\underbrace{0}_{0}\underbrace{01}_{1}\underbrace{011}_{2}\underbrace{0}_{0}\underbrace{011}_{2}\underbrace{01}_{1}\underbrace{0}_{0}\underbrace{01}_{1}\underbrace{011}_{2}\underbrace{01}_{1}\underbrace{0}_{0}\underbrace{011}_{2}\underbrace{0}_{0}\underbrace{01}_{1}\underbrace{011}_{2}\text{ etc.}$$

Si un carré apparaît dans le mot transformé, c'est qu'on peut trouver dans le mot de Morse un motif carré uu. Comme σ' est défini à partir de trois motifs commençant par 0 (« 011 », « 01 » et « 0 »), on en déduit que u commence par 0 et que le carré uu est suivi d'un 0 dans le mot de Morse infini. On aurait donc trouvé un motif de la forme $0v0v0$ dans le mot de Morse infini. Or nous avons vu que c'était impossible donc il n'y a pas de carré dans le mot de Morse transformé.

Conclusion

Suite à nos recherches, nous pouvons affirmer qu'il est possible de créer un mot infini à deux caractères évitant les cubes et un mot infini à trois caractères évitant les carrés.