Paperfolding and Words Avoiding Repetitions

Narad Rampersad

School of Computer Science
University of Waterloo

8 November 2006

(Joint work with Jui-Yi Kao, Jeffrey Shallit, and Manuel Silva)

Folding a Piece of Paper

- Take an ordinary 8.5×11 piece of paper and fold it in half.

Folding a Piece of Paper

- Take an ordinary 8.5×11 piece of paper and fold it in half.
- Now unfold the paper and record the pattern of hills and valleys created, writing 0 for a hill and 1 for a valley.

$$0$$

Folding a Piece of Paper

- Take an ordinary 8.5 $\times$ 11 piece of paper and fold it in half.
- Now unfold the paper and record the pattern of hills and valleys created, writing 0 for a hill and 1 for a valley.

$$0$$

- Now fold the paper twice, unfold, and record the pattern of hills and valleys.

$$0 \quad 0 \quad 1$$

Folding a Piece of Paper

- Take an ordinary 8.5×11 piece of paper and fold it in half.
- Now unfold the paper and record the pattern of hills and valleys created, writing 0 for a hill and 1 for a valley.

$$0$$

- Now fold the paper twice, unfold, and record the pattern of hills and valleys.

$$0 \quad 0 \quad 1$$

- Now fold three times, unfold, and record the pattern.

$$0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 1 \quad 1$$

Folding a Piece of Paper

- Take an ordinary 8.5 $\times$ 11 piece of paper and fold it in half.
- Now unfold the paper and record the pattern of hills and valleys created, writing 0 for a hill and 1 for a valley.

$$0$$

- Now fold the paper twice, unfold, and record the pattern of hills and valleys.

$$0 \quad 0 \quad 1$$

- Now fold three times, unfold, and record the pattern.

$$0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 1 \quad 1$$

- Now fold infinitely (!) many times. After unfolding, you get the following infinite sequence.

$$0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 1 \quad 1 \quad 0 \quad 0 \quad 0 \quad 1 \quad 1 \quad 0 \quad 1 \quad 1 \quad \cdots$$

Some Notation

For any word x over $\{0, 1\}$, let $\overline{x}$ denote the word obtained from x by changing 0's to 1's and 1's to 0's. Let x^R denote the reversal of x.

Example

If $x = 0111$, then $\overline{x} = 1000$, $x^R = 1110$, and $\overline{x}^R = 0001$.

Perturbed Symmetry

Definition

For $i \geq 0$, let $c_i \in \{0, 1\}$ and define the sequence of words

$$\begin{aligned} F_0 &= c_0 \\ F_1 &= F_0\, c_1\, \overline{F_0}^R \\ F_2 &= F_1\, c_2\, \overline{F_1}^R \\ &\vdots \end{aligned}$$

Then

$$\mathbf{f} = \lim_{i \to \infty} F_i$$

is a paperfolding word.

Perturbed Symmetry

Example

Taking $c_i = 0$ for all $i \geq 0$, one obtains the sequence

$$\begin{aligned} F_0 &= 0 \\ F_1 &= 0\,0\,1 \\ F_2 &= 001\,0\,011 \\ &\vdots \end{aligned}$$

which converges, in the limit, to the ordinary paperfolding word

$$001001100011011 0\cdots$$

A Recursive Definition

Definition

A paperfolding word $\mathbf{f} = f_0 f_1 f_2 \cdots$ over the alphabet $\{0, 1\}$ satisfies the following recursive definition: there exists $a \in \{0, 1\}$ such that

$$\begin{aligned} f_{4n} &= a, \quad n \geq 0 \\ f_{4n+2} &= \overline{a}, \quad n \geq 0 \\ (f_{2n+1})_{n\geq 0} & \quad \text{is a paperfolding word.} \end{aligned}$$

Definition

The ordinary paperfolding word

$$0010011000110110\cdots$$

is the paperfolding word uniquely characterized by $f_{2^m-1} = 0$ for all $m \geq 0$.

A Recursive Definition

Example

Consider the odd indexed terms of the ordinary paperfolding word:

$$\begin{array}{llcccccccccccccccc} \mathbf{f} & = & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & \cdots \\ \mathbf{f}' & = & & 0 & & 0 & & 1 & & 0 & & 0 & & 1 & & 1 & & \cdots \end{array}$$

Notice that $\mathbf{f} = \mathbf{f}'$.
The word $\mathbf{f}'$ will always be a paperfolding word for any $\mathbf{f}$, but in general one will not have $\mathbf{f} = \mathbf{f}'$.

The Toeplitz Construction

- Start with an infinite sequence of gaps, denoted ?.

? ? ? ? ? ? ? ? ? ? ? ? ? ? ? ...

- Fill every other gap with alternating 0's and 1's.

0 ? 1 ? 0 ? 1 ? 0 ? 1 ? 0 ? 1 ...

- Repeat.

0 0 1 ? 0 1 1 ? 0 0 1 ? 0 1 1 ...

0 0 1 0 0 1 1 ? 0 0 1 1 0 1 1 ...

0 0 1 0 0 1 1 0 0 0 1 1 0 1 1 ...

The Toeplitz Construction

- In the limit, one again obtains the ordinary paperfolding word

$$0010011000110110\cdots$$

- At each step, one may choose to fill in the gaps by either

$$0101010101\cdots$$

  or

$$1010101010\cdots .$$

- Different choices at each step results in the construction of different paperfolding words.
- Words constructed by such a process are called Toeplitz words.

Structure in the Paperfolding Words

Theorem

No paperfolding word is ultimately periodic.

Theorem (Allouche 1992)

Let $\mathbf{f}$ *be a paperfolding word. Any subword of* $\mathbf{f}$ *of length at least* 7 *cannot occur at both an odd and an even position of* $\mathbf{f}$*.*

Theorem (Allouche 1992)

A paperfolding word has exactly $4n$ *distinct subwords of length* n *for* $n \geq 7$*.*

Repetitions in Words

Definition

A square (or 2-power) is a non-empty word of the form ww (or w^2). A word is squarefree if none of its subwords are squares.

Definition

A cube (or 3-power) is a non-empty word of the form www (or w^3). A word is cubefree if none of its subwords are cubes.

Example

- *tumtum* (as in “So rested he by the Tumtum tree”) is a square.
- *hohoho* is a cube.

Repetitions in Words

Definition

A overlap (or 2^+-power) is a non-empty word of the form *axaxa*, where *a* is a letter and *x* is a (possibly empty) word.

Definition

A 3^+-power is a non-empty word of the form *axaxaxa*, where *a* is a letter and *x* is a (possibly empty) word.

Example

- *entente* is an overlap (*chevauchement* en français).
- 0110110110 is a 3^+-power.

One generalizes these definitions to k-powers and k^+-powers in the obvious way.

Avoiding Repetitions in Words

Theorem (Thue 1906)

There exists an infinite squarefree word

$$\mathbf{w} = 210201210120210\cdots$$

over the alphabet $\{0, 1, 2\}$.

Proof (sketch).

The word $\mathbf{w}$ is obtained by iterating the map $2 \to 210$, $1 \to 20$, $0 \to 1$:

$$2 \to 210 \to 210201 \to 210201210120 \to \cdots$$

□

Avoiding Repetitions in Words

Theorem (Thue 1912)

There exists an infinite overlapfree word

$$\mathbf{t} = 0110100110010110\cdots$$

over the alphabet $\{0, 1\}$.

Proof (sketch).

The word $\mathbf{t}$ is obtained by iterating the map $0 \to 01$, $1 \to 10$:

$$0 \to 01 \to 0110 \to 01101001 \to \cdots$$

□

Avoiding Large Repetitions in Words

Can we avoid all sufficiently large squares over a binary alphabet?

Theorem (Entringer, Jackson, and Schatz 1974)

There exists an infinite binary word $\mathbf{x}$ *containing no squares* xx *where* $|x| \geq 3$.

Proof (sketch).

Let $\mathbf{w}$ be any infinite squarefree word over $\{0, 1, 2\}$. Apply the map $0 \rightarrow 1010$, $1 \rightarrow 1100$, $2 \rightarrow 0111$ to $\mathbf{w}$ to obtain $\mathbf{x}$; e.g. if

$$\mathbf{w} = 210201210120210\cdots,$$

then

$$\mathbf{x} = 011111001010011110 10\cdots.$$

□

Repetitions in Paperfolding Words

Theorem (Prodinger and Urbanek 1979)

For the ordinary paperfolding word $\mathbf{f}$*, if* ww *is a non-empty subword of* $\mathbf{f}$*, then* $|w| \in \{1, 3, 5\}$*.*

Theorem (Allouche and Bousquet-Mélou 1994)

For any paperfolding word $\mathbf{f}$*, if* ww *is a non-empty subword of* $\mathbf{f}$*, then* $|w| \in \{1, 3, 5\}$*.*

Corollary (Allouche and Bousquet-Mélou 1994)

For any paperfolding word $\mathbf{f}$*,* $\mathbf{f}$ *contains no* 4*-powers and no cubes except* 000 *and* 111*. In particular,* $\mathbf{f}$ *contains no* 3^+*-power.*

A Language-theoretic Consequence

Corollary (Lehr 1992; Allouche and Bousqet-Mélou 1994)

The language consisting of all subwords of paperfolding words is not context-free.

Proof.

It is clear from the pumping lemma that any infinite context-free language contains words with arbitrarily large repetitions. □

Arithmetic Subsequences

Definition

Let

$$\mathbf{w} = w_0 w_1 w_2 \cdots$$

be a word. An arithmetic subsequence of difference j is a word

$$w_i w_{i+j} w_{i+2j} \cdots w_{i+tj}$$

for some i, t.

Example

If

$$\mathbf{w} = w_0 w_1 w_2 \cdots = 0110100110010110 \cdots$$

then an arithmetic subsequence of difference 3 of $\mathbf{w}$ is

$$w_1 w_4 w_7 w_{10} = 1110.$$

van der Waerden's Theorem

- Recall, a word is squarefree if no subword is a square.
- Does there exists an infinite word such that no arithmetic subsequence is a square?
- Clearly, no. What about trying to avoid cubes, or 4-powers, etc.?

Theorem (van der Waerden 1927)

For any infinite word $\mathbf{w}$ *over a finite alphabet* A*, there exists* $a \in A$ *such that for all* $m \geq 1$*,* $\mathbf{w}$ *contains* a^m *in arithmetic progressions.*

- Suppose we only try to avoid repetitions in certain types of arithmetic progressions: e.g. arithmetic progressions of odd difference.

Subsequences of the Paperfolding Words

Theorem (Avgustinovich, Fon-Der-Flaass, and Frid 2003)

If w is a finite arithmetic subsequence of odd difference of a paperfolding word, then w is a subword of a paperfolding word.

Example

Take the first 15 symbols of the ordinary paperfolding word:

$$f_0 f_1 \cdots f_{14} = 001001100011011.$$

Then

$$\begin{aligned} f_0 f_3 \cdots f_{12} &= 00100 \\ f_1 f_4 \cdots f_{13} &= 00011 \\ f_2 f_5 \cdots f_{14} &= 11011. \end{aligned}$$

Subsequences of the Paperfolding Words

Example

Continuing, if

$$f_0 f_1 \cdots f_{14} = 001001100011011,$$

then

$$\begin{aligned} f_0 f_5 f_{10} &= 011 \\ f_1 f_6 f_{11} &= 011 \\ f_2 f_7 f_{12} &= 100 \\ f_3 f_8 f_{13} &= 001 \\ f_4 f_9 f_{14} &= 001. \end{aligned}$$

One verifies that each of these are subwords of **f**.

Subsequences of the Paperfolding Words

- Recall that every paperfolding word is 3^+-powerfree.

Corollary

There exists an infinite word over a binary alphabet that contains no 3^+-powers in arithmetic progressions of odd difference. Indeed, all paperfolding words have this property.

- The 3^+ above is optimal; it cannot be replaced by 3.
- If we increase the alphabet size, can we avoid squares in all arithmetic progressions of odd difference?

Repetitions in Arithmetic Progressions

Theorem (Carpi 1988)

There exists an infinite word over a four letter alphabet that avoids squares in arithmetic progressions of odd difference.

Let $\mathbf{f} = f_0 f_1 f_2 \cdots$ be any paperfolding word over $\{1, 4\}$. Define $\mathbf{v} = v_0 v_1 v_2 \cdots$ by

$$\begin{aligned} v_{4n} &= 2 \\ v_{4n+2} &= 3 \\ v_{2n+1} &= f_{2n+1}, \end{aligned}$$

for all $n \geq 0$.

In other words, we have recoded the periodic subsequence formed by taking the even positions of $\mathbf{f}$ by mapping $1 \to 2$ and $4 \to 3$ (or vice-versa).

Proof of Carpi's Theorem

Example

If

$$\mathbf{f} = 1141144111441441\cdots$$

is the ordinary paperfolding word over $\{1, 4\}$, then

$$\mathbf{v} = 2131243121342431\cdots.$$

Theorem

Let $\mathbf{v}$ *be any word obtained from a paperfolding word* $\mathbf{f}$ *by the construction described above. Then the word* $\mathbf{v}$ *contains no squares in any arithmetic progression of odd difference.*

Proof of Carpi's Theorem

By the construction of **v**, any arithmetic subsequence

$$w = v_{i_0} v_{i_1} \cdots v_{i_k}$$

of odd difference of **v** can be obtained from the corresponding subsequence

$$x = f_{i_0} f_{i_1} \cdots f_{i_k}$$

of **f** by recoding the symbols in either the even positions of x or the odd positions of x by mapping $1 \to 2$ and $4 \to 3$ (or vice-versa).
Note that this recoding cannot create any new squares.

Proof of Carpi's Theorem

Now suppose that **v** contains a square *ww* in an arithmetic progression of odd difference. Let *xx* be the corresponding subsequence of **f**.
By previous results, $|x| \in \{1, 3, 5\}$ and hence $|w| \in \{1, 3, 5\}$.
Clearly, $|w| = 1$ is impossible.
If $|w| = 3$, then *ww* has one of the forms

$$(*2*)(3*2)$$

$$(*3*)(2*3)$$

$$(2*3)(*2*)$$

$$(3*2)(*3*)$$

where the $*$ denotes an arbitrary symbol from $\{1, 4\}$.
Clearly, none of these can be squares.
A similar argument applies for $|w| = 5$.

Avoiding Overlaps in Odd Difference A.P.

Theorem

There exists an infinite word over a ternary alphabet that contains no 2^+-powers (overlaps) and no squares xx, $|x| \geq 2$, in arithmetic progressions of odd difference.

Proof (sketch).

Let $\mathbf{v}$ be any word obtained from a paperfolding word by the construction described above. Let h map $1 \to 00$, $2 \to 11$, $3 \to 12$, $4 \to 02$. Then $\mathbf{w} = h(\mathbf{v})$ has the desired properties; e.g. if

$$\mathbf{v} = 2131243121342431 \cdots ,$$

then

$$\mathbf{w} = 110012001102120011 \cdots .$$

□

Avoiding Large Squares in Odd Difference A.P.

Theorem

There exists an infinite word over a binary alphabet that contains no squares xx with $|x| \geq 3$ in any arithmetic progression of odd difference.

Proof (sketch).

Let $\mathbf{v}$ be any word obtained from a paperfolding word by the construction described above. Let h map $1 \to 0110$, $2 \to 0101$, $3 \to 0001$, $4 \to 0111$. Then $\mathbf{w} = h(\mathbf{v})$ has the desired properties; e.g. if

$$\mathbf{v} = 2131243121342431 \cdots ,$$

then

$$\mathbf{w} = 0101011000010110010 1 \cdots .$$

□

Higher Dimensions

An infinite word over a finite alphabet A is a map $\mathbf{w}$ from $\mathbb{N}$ to A, where we write w_n for $\mathbf{w}(n)$. This inspires the following generalization.

Definition

A 2-dimensional word is a map $\mathbf{w}$ from $\mathbb{N}^2$ to A, where we write $w_{m,n}$ for $\mathbf{w}(m, n)$.

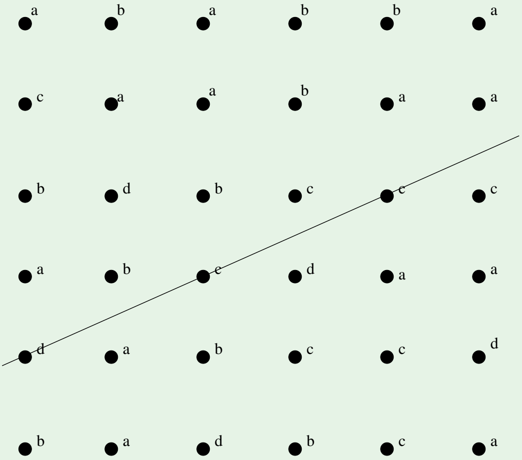
Definition

A word $\mathbf{x}$ is a line of $\mathbf{w}$ if there exists i_1, i_2, j_1, j_2, such that $\gcd(j_1, j_2) = 1$ and for $t \geq 0$

$$x_t = w_{i_1+j_1 t, i_2+j_2 t}.$$

Higher Dimensions

Example



Here $\mathbf{x} = \mathrm{dcc}\cdots$ is a line.

Higher Dimensions

Theorem (Carpi 1988)

There exists a 2-dimensional word $\mathbf{w}$ *over a 16-letter alphabet, such that every line of* $\mathbf{w}$ *is squarefree.*

Proof of Carpi's 2D construction

Proof.

Let $\mathbf{u} = u_0 u_1 u_2 \cdots$ and $\mathbf{v} = v_0 v_1 v_2 \cdots$ be any infinite words over the alphabet $A = \{1, 2, 3, 4\}$ that avoid squares in all arithmetic progressions of odd difference. We define $\mathbf{w}$ over the alphabet $A \times A$ by

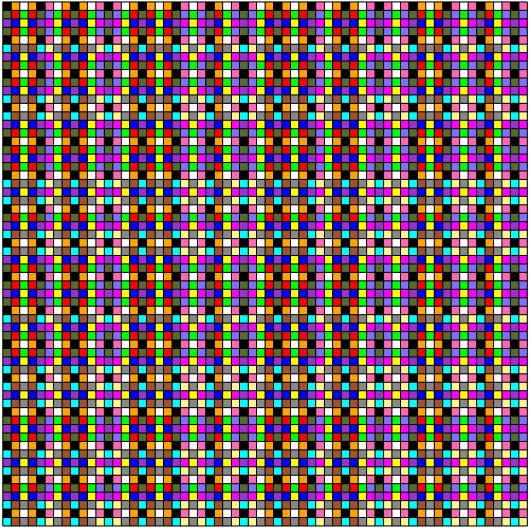
$$w_{m,n} = (u_m, v_n).$$

Consider an arbitrary line

$$\begin{aligned} \mathbf{x} &= (w_{i_1+j_1 t, i_2+j_2 t})_{t \geq 0}, \\ &= (u_{i_1+j_1 t}, v_{i_2+j_2 t})_{t \geq 0}, \end{aligned}$$

for some i_1, i_2, j_1, j_2, with $\gcd(j_1, j_2) = 1$. Without loss of generality, we may assume j_1 is odd. Then the word $(u_{i_1+j_1 t})_{t \geq 0}$ is an arithmetic subsequence of odd difference of $\mathbf{u}$ and hence is squarefree. The line $\mathbf{x}$ is therefore also squarefree. □

Tiling Based on Carpi's 2D Word



Avoiding 3^+-powers on the Integer Lattice

Theorem

There exists a 2-dimensional word $\mathbf{w}$ *over a 4-letter alphabet, such that every line of* $\mathbf{w}$ *is* 3^+*-power-free.*

Proof.

Let $\mathbf{u} = u_0 u_1 u_2 \cdots$ and $\mathbf{v} = v_0 v_1 v_2 \cdots$ be any paperfolding words. Then $\mathbf{u}$ and $\mathbf{v}$ each avoid 3^+-powers in all arithmetic progressions of odd difference. We now define $\mathbf{w}$ by

$$w_{m,n} = (u_m, v_n).$$

□

Avoiding Overlaps on the Integer Lattice

Theorem

There exists a 2-dimensional word $\mathbf{w}$ *over a 9-letter alphabet, such that every line of* $\mathbf{w}$ *is* 2^+*-power-free (overlapfree).*

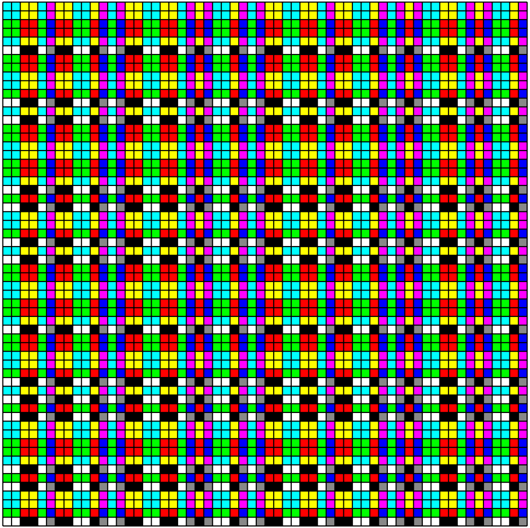
Proof.

Let $\mathbf{u} = u_0 u_1 u_2 \cdots$ and $\mathbf{v} = v_0 v_1 v_2 \cdots$ be any words over $\{0, 1, 2\}$ that avoid overlaps in all arithmetic progressions of odd difference. We now define $\mathbf{w}$ by

$$w_{m,n} = (u_m, v_n).$$

□

Tiling Based on 2D Overlapfree Word



Avoiding Large Squares on the Integer Lattice

Theorem

There exists a 2-dimensional word $\mathbf{w}$ *over a 4-letter alphabet, such that every line of* $\mathbf{w}$ *avoids squares* xx*, where* $|x| \geq 3$*.*

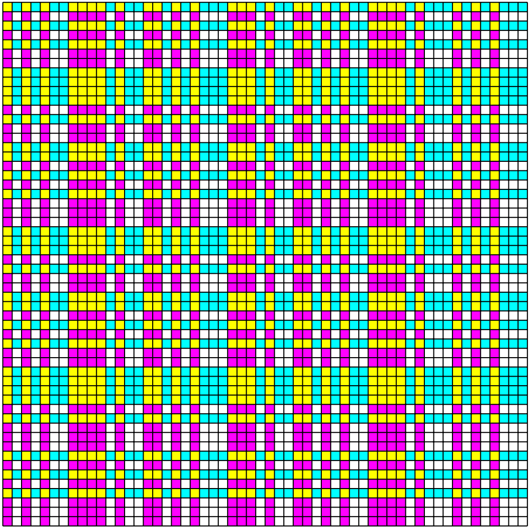
Proof.

Let $\mathbf{u} = u_0 u_1 u_2 \cdots$ and $\mathbf{v} = v_0 v_1 v_2 \cdots$ be any words over $\{0, 1\}$ that avoid squares xx, $|x| \geq 3$, in all arithmetic progressions of odd difference. We now define $\mathbf{w}$ by

$$w_{m,n} = (u_m, v_n).$$

□

Tiling that Avoids Large Squares



Open Problems

- Recall that the language of all subwords of paperfolding words is not context-free. What about its complement? It is known that if the complement is context-free, it must be inherently ambiguous.
- Find the optimal alphabet sizes for the 2D constructions described above.
- We have only discussed here arithmetic progressions of odd difference. What about other differences? Carpi's 1988 paper has some additional results in this regard.