Minimal Overlapping & Exact Matches in Words

Brian Miceli
Department of Mathematics, Trinity University
Permutation Patterns 2011
Cal Poly, SLO

June 24^{th}, 2011

Joint work with Jeff Remmel

A Quick Outline

- Some Quick Motivation
- What is the Problem?
- Some Mathematics

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

Definitions

Definition: For any alphabet, $\mathcal{A}$, we let $\mathcal{A}^*$ denote the set of all words over $\mathcal{A}$. We denote the *empty word* by ϵ and we let $\mathcal{A}^+ = \mathcal{A}^* - \{\epsilon\}$.

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

Definitions

Definition: For any alphabet, $\mathcal{A}$, we let $\mathcal{A}^*$ denote the set of all words over $\mathcal{A}$. We denote the *empty word* by ϵ and we let $\mathcal{A}^+ = \mathcal{A}^* - \{\epsilon\}$.

Note: For this talk, we will concern ourselves with the alphabets $\mathbb{N} = \{1, 2, 3, \ldots\}$, $\mathbb{N}_0 = \{0, 1, 2, 3, \ldots\}$, and $\mathbb{B} = \{0, 1\}$.

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

TRINITY UNIVERSITY

Definitions

Given a word $w = w_1 w_2 \cdots w_n \in \mathbb{N}^*$, we give the following definitions.

Definition: The *weight* of w is $\Sigma w = w_1 + w_2 + \cdots + w_n$.

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

Definitions

Given a word $w = w_1 w_2 \cdots w_n \in \mathbb{N}^*$, we give the following definitions.

Definition: The *weight* of w is $\Sigma w = w_1 + w_2 + \cdots + w_n$.

Definition: The *length* of w is $|w| = n$.

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

Definitions

Given a word $w = w_1 w_2 \cdots w_n \in \mathbb{N}^*$, we give the following definitions.

Definition: The *weight* of w is $\Sigma w = w_1 + w_2 + \cdots + w_n$.

Definition: The *length* of w is $|w| = n$.

Definition: $z(w) = \prod_{i=1}^{n} z_{w_i}$.

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

Definitions

Given a word $w = w_1 w_2 \cdots w_n \in \mathbb{N}^*$, we give the following definitions.

Definition: The *weight* of w is $\Sigma w = w_1 + w_2 + \cdots + w_n$.

Definition: The *length* of w is $|w| = n$.

Definition: $z(w) = \prod_{i=1}^{n} z_{w_i}$.

Definition: $n_a(w)$ is the number of occurrences of a in w.

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

Motivation

We begin with a basic question to look at:

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

Motivation

We begin with a basic question to look at:

How many words in $\mathbb{N}^*$ of weight n contain exactly one occurrence of the word 13?

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

Examples to Gather Data

We first generate examples of words with exactly one containment of 13.

Example: There are 0 words of weight $n = 1, 2, 3$.

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

Examples to Gather Data

We first generate examples of words with exactly one containment of 13.

Example: There are 0 words of weight $n = 1, 2, 3$.

Example: $(n = 4)$ There is 1 word - 13.

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

Examples to Gather Data

We first generate examples of words with exactly one containment of 13.

Example: There are 0 words of weight $n = 1, 2, 3$.

Example: $(n = 4)$ There is 1 word - 13.

Example: $(n = 5)$ There are 4 words - 14, 23, 1$\underline{13}$, $\underline{13}$1.

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

Examples to Gather Data

We first generate examples of words with exactly one containment of 13.

Example: There are 0 words of weight $n = 1, 2, 3$.

Example: $(n = 4)$ There is 1 word - 13.

Example: $(n = 5)$ There are 4 words - 14, 23, $1\underline{13}$, $\underline{13}1$.

Example: $(n = 6)$ There are 12 words - 15, 24, $1\underline{14}$, $\underline{14}1$, 33, $1\underline{23}$, $\underline{13}2$, $2\underline{13}$, $\underline{23}1$, $11\underline{13}$, $1\underline{13}1$, $\underline{13}11$.

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

Examples to Gather Data

We first generate examples of words with exactly one containment of 13.

Example: There are 0 words of weight $n = 1, 2, 3$.

Example: $(n = 4)$ There is 1 word - 13.

Example: $(n = 5)$ There are 4 words - 14, 23, $1\underline{13}$, $\underline{13}1$.

Example: $(n = 6)$ There are 12 words - 15, 24, $1\underline{14}$, $\underline{14}1$, 33, $1\underline{23}$, $\underline{13}2$, $2\underline{13}$, $\underline{23}1$, $11\underline{13}$, $1\underline{13}1$, $\underline{13}11$.

Example: $(n = 7, 8)$ There are 30 words of weight 7 and 68 of weight 8.

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

Now We Have Some Data

We have a sequence of terms that looks like:

$$0, 0, 0, 1, 4, 12, 30, 68, \ldots.$$

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

Now We Have Some Data

We have a sequence of terms that looks like:

$$0, 0, 0, 1, 4, 12, 30, 68, \ldots.$$

The OEIS says that this sequence may be the number of binary words of length n that contain exactly one occurrence of the binary subword 001.

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

Now We Have Some Data

We have a sequence of terms that looks like:

$$0, 0, 0, 1, 4, 12, 30, 68, \ldots .$$

The OEIS says that this sequence may be the number of binary words of length n that contain exactly one occurrence of the binary subword 001.

This makes sense.

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

The Problem

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

The Problem

Suppose we have the word $b = 0^{k_1}10^{k_2}1\cdots0^{k_m}10^{k_{m+1}} \in B^*$ where $k_i \geq 0$, $m \geq 2$, and $min(k_1, k_{m+1}) > max(k_2, \ldots, k_m)$.

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

The Problem

Suppose we have the word $b = 0^{k_1}10^{k_2}1\cdots0^{k_m}10^{k_{m+1}} \in B^*$ where $k_i \geq 0$, $m \geq 2$, and $min(k_1, k_{m+1}) > max(k_2, \ldots, k_m)$.

Can we find a generating function for the number of consecutive occurrences of b over the set of words from the alphabet $\mathbb{B}$?

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

More Definitions

Let $w = w_1w_2\cdots w_n \in \mathbb{N}^n$ and $u = u_1u_2\cdots u_j \in \mathbb{N}^j$ with $j \leq n$.

Definition: w has an *exact u-match starting at position i* if

$$w_iw_{i+1}\cdots w_{i+j-1} = u.$$

More Definitions

Let $w = w_1 w_2 \cdots w_n \in \mathbb{N}^n$ and $u = u_1 u_2 \cdots u_j \in \mathbb{N}^j$ with $j \leq n$.

Definition: w has an *exact u-match starting at position i* if $w_i w_{i+1} \cdots w_{i+j-1} = u$.

Definition: w has an *endpoint embedding of u starting at position i* if $w_i \geq u_1$, $w_{i+j-1} \geq u_j$, and $w_{i+k} = u_{k+1}$ for $2 \leq k \leq j-1$.

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

More Definitions

Let $w = w_1 w_2 \cdots w_n \in \mathbb{N}^n$ and $u = u_1 u_2 \cdots u_j \in \mathbb{N}^j$ with $j \leq n$.

Definition: w has an *exact u-match starting at position i* if $w_i w_{i+1} \cdots w_{i+j-1} = u$.

Definition: w has an *endpoint embedding of u starting at position i* if $w_i \geq u_1$, $w_{i+j-1} \geq u_j$, and $w_{i+k} = u_{k+1}$ for $2 \leq k \leq j-1$.

Definition: $ex_u(w)$ is the number of exact u-matches in w, and $endp_u(w)$ is the number of endpoint embedding of u in w.

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

An Example

Example: Let $u = 312$ and $w = 2413113122312$.

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

An Example

Example: Let $u = 312$ and $w = 2413113122312$.

Then the word w has exact u matches at starting positions 7 and 11: $w = 241311\underline{312}2\underline{312}$.

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

An Example

Example: Let $u = 312$ and $w = 2413113122312$.

Then the word w has exact u matches at starting positions 7 and 11: $w = 241311\underline{312}2\underline{312}$.

The word w also has endpoint embeddings of u at starting positions 2, 7, and 11: $w = 2\underline{413}11\underline{312}2\underline{312}$.

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

An Example

Example: Let $u = 312$ and $w = 2413113122312$.

Then the word w has exact u matches at starting positions 7 and 11: $w = 241311\underline{312}2\underline{312}$.

The word w also has endpoint embeddings of u at starting positions 2, 7, and 11: $w = 2\underline{413}11\underline{312}2\underline{312}$.

So, $ex_u(w) = 2$ and $endp_u(w) = 3$.

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

More Definitions

Definition: $u \in \mathbb{N}^j$ has the *end point minimal overlapping property* if $i = n(j-1)+1$ is the smallest i such that there exists $w \in \mathbb{N}^i$ with $ep_u(w) = n$. We denote the set of such w's by $\mathcal{EPMP}_{u,n(j-1)+1}$, and we refer to the elements of this set as *end point embedding maximum packings of u.*

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

More Definitions

Definition: $epmp_{u,n(j-1)+1} = |\mathcal{EPMP}_{u,n(j-1)+1}|$.

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

More Definitions

Definition: $epmp_{u,n(j-1)+1} = |\mathcal{EPMP}_{u,n(j-1)+1}|$.

Definition: $$epmp_{u,n(j-1)+1}(z) = \sum_{w \in \mathcal{EPMP}_{u,n(j-1)+1}} z^{\Sigma w}.$$

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

More Definitions

Definition: $epmp_{u,n(j-1)+1} = |\mathcal{EPMP}_{u,n(j-1)+1}|$.

Definition: $epmp_{u,n(j-1)+1}(z) = \sum_{w \in \mathcal{EPMP}_{u,n(j-1)+1}} z^{\Sigma w}$.

Definition: $epmp_{u,n(j-1)+1}(z_0, z_1, \ldots) = \sum_{w \in \mathcal{EPMP}_{u,n(j-1)+1}} z(w)$.

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

A Computational Example

Example: Suppose $u = 5324$ and $w \in \mathcal{EPMP}_{u,3n+1}$.

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

A Computational Example

Example: Suppose $u = 5324$ and $w \in \mathcal{EPMP}_{u,3n+1}$. Then $w = w_1 32 w_4 32 \ldots w_{3n-2} 32 w_{3n+1}$, where $w_1 \geq 5$, $w_{3n+1} \geq 4$, and $w_{3k+1} \geq max(4,5)$ for $1 \leq k \leq n-1$. So, $epmp_{5324,3n+1}(z_0, z_1, \ldots) =$

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

A Computational Example

Example: Suppose $u = 5324$ and $w \in \mathcal{EPMP}_{u,3n+1}$. Then $w = w_1 32 w_4 32 \ldots w_{3n-2} 32 w_{3n+1}$, where $w_1 \geq 5$, $w_{3n+1} \geq 4$, and $w_{3k+1} \geq max(4,5)$ for $1 \leq k \leq n-1$. So,

$epmp_{5324,3n+1}(z_0, z_1, \ldots) =$

$$\left(\sum_{i \geq 5} z_i\right)\left(\sum_{i \geq 4} z_i\right)\left(\sum_{i \geq max(4,5)} z_i\right)^{n-1} (z_2 z_3)^n.$$

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

TRINITY UNIVERSITY

A Main Theorem

Theorem: Suppose $u \in \mathbb{N}^j$ has the end point minimal embedding property. Then

$$\sum_{n \geq 0} t^n \sum_{w \in \mathbb{N}^n} x^{ep_u(w)} z(w) =$$

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

A Main Theorem

Theorem: Suppose $u \in \mathbb{N}^j$ has the end point minimal embedding property. Then

$$\sum_{n\geq 0} t^n \sum_{w\in\mathbb{N}^n} x^{ep_u(w)} z(w) = \frac{1}{1-((\sum_{i\geq 0} z_i)t + \sum_{n\geq 1} t^{n(j-1)+1}(x-1)^n epmp_{u,n(j-1)+1}(z_0,z_1,\ldots))}.$$

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

A Main Theorem

Theorem: Suppose $u \in \mathbb{N}^j$ has the end point minimal embedding property. Then

$$\sum_{n \geq 0} t^n \sum_{w \in \mathbb{N}^n} x^{ep_u(w)} z(w) =$$

$$\frac{1}{1 - ((\sum_{i \geq 0} z_i)t + \sum_{n \geq 1} t^{n(j-1)+1} (x-1)^n epmp_{u,n(j-1)+1}(z_0, z_1, \ldots))}.$$

Proof: Brick tabloids, involution principle, and lots of cases.

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

More Definitions

Definition: $u \in \mathbb{N}^j$ has the *exact match minimal overlapping property* if $i = n(j-1)+1$ is the smallest i such that there exists $w \in \mathbb{N}^i$ with $ex_u(w) = n$. We denote the set of such w's by $\mathcal{EMP}_{u,n(j-1)+1}$, and we refer to the elements of this set as *exact match maximum packings of* u.

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

More Definitions

Definition: $u \in \mathbb{N}^j$ has the *exact match minimal overlapping property* if $i = n(j-1)+1$ is the smallest i such that there exists $w \in \mathbb{N}^i$ with $ex_u(w) = n$. We denote the set of such w's by $\mathcal{EMP}_{u,n(j-1)+1}$, and we refer to the elements of this set as *exact match maximum packings of* u.

Example: For any $k \in \mathbb{N}$, $u = 01^k0 \in \mathcal{EMP}_{u,n(k+1)+1}$.

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

Another Main Result

Theorem: Suppose $u = u_1 u_2 \cdots u_j \in \mathbb{N}_0^j$ with $j \geq 3$ with $min(u_1, u_j) > max(u_2, \ldots, u_{j-1})$.

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

Another Main Result

Theorem: Suppose $u = u_1 u_2 \cdots u_j \in \mathbb{N}_0^j$ with $j \geq 3$ with $min(u_1, u_j) > max(u_2, \ldots, u_{j-1})$. If $\Theta(u) = 0^{u_1} 1 0^{u_2} 1 \cdots 1 0^{u_j}$, then

$$\sum_{v \in \mathbb{B}^*} x^{ex_{\Theta(u)}(v)} t^{n_1(v)} z^{|v|} = \text{Some big, messy generating function.}$$

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

An Application

Example: Let $u = 312$, so that $\Theta(u) = 0^3 1010^2$, which does not have the minimal overlap property. Then we can use our previous theorem to find the following:

$$\sum_{\substack{v \in \mathbb{B}^* \\ ex_{0^3 1010^2}(v)=0}} z^{|v|} =$$

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

An Application

Example: Let $u = 312$, so that $\Theta(u) = 0^3 1010^2$, which does not have the minimal overlap property. Then we can use our previous theorem to find the following:

$$\sum_{\substack{v \in \mathbb{B}^* \\ ex_{0^3 1010^2}(v)=0}} z^{|v|} = \frac{1 + z^6 + z^7}{1 - 2z + z^6 - z^7 - z^8}.$$

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

Future Directions

Question: Is there anything special about binary words?

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

Future Directions

Question: Is there anything special about binary words?

Answer: Yes & No. (This is joint work with Tom Langley)

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli

TRINITY UNIVERSITY

The End

Thanks to the Cal Poly Math Department and all of the organizers.

Brian Miceli
bmiceli@trinity.edu
http://www.trinity.edu/bmiceli