Strict Bounds for Pattern Avoidance

F. Blanchet-Sadri [1] Brent Woodhouse [2]

[1]University of North Carolina at Greensboro

[2]Purdue University

To be presented at DLT 2013

This material is based upon work supported by the National Science Foundation under Grant No. DMS–1060775.

Outline

1. Introduction
2. Two sequences of unavoidable patterns
3. The power series approach
4. Derivation of the strict bounds
5. Extension to partial words
6. Conclusion

1. Introduction

- Cassaigne conjectured in 1994 that any pattern with m distinct variables of length at least $3(2^{m-1})$ is avoidable over 2 letters, and any pattern with m distinct variables of length at least 2^m is avoidable over 3 letters.
- Building upon the work of Rampersad and the power series techniques of Bell and Goh, we obtain both of these suggested strict bounds.
- Similar bounds are also obtained for pattern avoidance in partial words, sequences where some characters are unknown.

Let Σ be an alphabet of letters, denoted by $a, b, c, \ldots$, and Δ be an alphabet of variables, denoted by $A, B, C, \ldots$.

- A pattern p is a word over Δ.
- A word w over Σ is an instance of p if there exists a non-erasing morphism $\varphi : \Delta^* \to \Sigma^*$ such that $\varphi(p) = w$.
- A word w is said to avoid p if no factor of w is an instance of p.

$\underline{aa}\ \underline{b}\ \underline{aa}\ c$ contains an instance of ABA while $abaca$ avoids AA

Avoidability and k-avoidability

- A pattern p is avoidable if there exist infinitely many words w over a finite alphabet such that w avoids p, or equivalently, if there exists an infinite word that avoids p.
- If p is avoided by infinitely many words over k letters, p is k-avoidable.
- If p is avoidable, the minimum k such that p is k-avoidable is called the avoidability index of p.

ABA is unavoidable while AA has avoidability index 3

- If a pattern p occurs in a pattern q, we say p divides q.

  $p = ABA$ divides $q = \underline{ABC}\ \underline{BB}\ \underline{ABC}\ A$, since we can map A to ABC and B to BB and this maps p to a factor of q

- If p divides q and p is k-avoidable, there exists an infinite word w over k letters that avoids p; w must also avoid q, thus q is necessarily k-avoidable. It follows that

  the avoidability index of $q \leq$ the avoidability index of p

- It is not known if it is generally decidable, given a pattern p and integer k, whether p is k-avoidable.
- Thus various authors compute avoidability indices and try to find bounds on them.
- Cassaigne's 1994 Ph.D. Thesis listed avoidability indices for unary, binary, and most ternary patterns (Ochem 2006 determined the remaining few avoidability indices for ternary patterns).
- Based on this data, Cassaigne conjectured in his thesis:
  - Any pattern with m distinct variables of length at least $3(2^{m-1})$ is avoidable over 2 letters;
  - Any pattern with m distinct variables of length at least 2^m is avoidable over 3 letters.
- Our main result is the affirmative answer to this long-standing conjecture of Cassaigne.

2. Two sequences of unavoidable patterns

Both bounds suggested by Cassaigne are strict.

Proposition

Let p be a k-unavoidable pattern over Δ and $A \in \Delta$ be a variable that does not occur in p. Then the pattern pAp is k-unavoidable.

Sequences of patterns that meet the bounds

Let $A_1, A_2, \ldots$ be distinct variables in Δ.

- $Z_0 = \varepsilon$ and for all $m \geq 0$, $Z_{m+1} = Z_m A_{m+1} Z_m$
  Since ε is k-unavoidable for every positive integer k, the previous proposition implies Z_m is k-unavoidable for all $m \in \mathbb{N}$ by induction on m. Thus Z_m is a 3-unavoidable pattern over m variables with length $2^m - 1$ for all $m \in \mathbb{N}$.

- $R_1 = A_1 A_1$ and for all $m \geq 1$, $R_{m+1} = R_m A_{m+1} R_m$
  Since $A_1 A_1$ is 2-unavoidable, the previous proposition implies R_m is 2-unavoidable for all $m \in \mathbb{N}$ by induction on m. Thus R_m is a 2-unavoidable pattern over m variables with length $3(2^{m-1}) - 1$ for all $m \in \mathbb{N}$.

3. The power series approach

Theorem

Let S be a set of words over k letters with each word of length at least two. Suppose that for each $i \geq 2$, the set S contains at most c_i words of length i. If the power series expansion of

$$B(x) := \left(1 - kx + \sum_{i \geq 2} c_i x^i\right)^{-1}$$

has non-negative coefficients, then there are at least $[x^n]B(x)$ words of length n over k letters that have no factors in S.

To count the number of words of length n avoiding a pattern p, we let S consist of all instances of p.

Rampersad, N.: Further applications of a power series method for pattern avoidance. *The Electronic Journal of Combinatorics* **18** (2011) P134

Bell and Goh's lemma (a useful upper bound)

Let $m \geq 1$ be an integer and p be a pattern over an alphabet $\Delta = \{A_1, \ldots, A_m\}$. Suppose that for $1 \leq i \leq m$, the variable A_i occurs $d_i \geq 1$ times in p. Let $k \geq 2$ be an integer and let Σ be a k-letter alphabet. Then for $n \geq 1$, the number of words of length n over Σ that are instances of the pattern p is no more than $[x^n]C(x)$, where

$$C(x) := \sum_{i_1 \geq 1} \cdots \sum_{i_m \geq 1} k^{i_1 + \cdots + i_m} x^{d_1 i_1 + \cdots + d_m i_m}$$

Note that this approach for counting instances of a pattern is based on the frequencies of each variable in the pattern, so it will not distinguish *AABB* and *ABAB*, for example.

Bell, J., Goh, T.L.: Exponential lower bounds for the number of words of uniform length avoiding a pattern. *Information and Computation* **205** (2007) 1295–1306

4. Derivation of the strict bounds

Lemma

Suppose $k \geq 2$ and $m \geq 1$ are integers and $\lambda > \sqrt{k}$. For any integer P and integers d_j for $1 \leq j \leq m$ such that $d_j \geq 2$ and $P = d_1 + \cdots + d_m$,

$$\prod_{i=1}^{m} \frac{1}{\lambda^{d_i} - k} \leq \left(\frac{1}{\lambda^2 - k}\right)^{m-1} \left(\frac{1}{\lambda^{P-2(m-1)} - k}\right)$$

Proof

The proof is by induction on m.

- For $m = 1$, $d_1 = P$ and the inequality is trivially satisfied.
- Suppose the inequality holds for m and $d_1 + d_2 + \cdots + d_{m+1} = P$ with $d_j \geq 2$ for $1 \leq j \leq m + 1$.
- Letting $P' = P - d_{m+1} = d_1 + \cdots + d_m$, the inductive hypothesis implies

$$\prod_{i=1}^{m} \frac{1}{\lambda^{d_i} - k} \leq \left(\frac{1}{\lambda^2 - k}\right)^{m-1} \left(\frac{1}{\lambda^{P' - 2(m-1)} - k}\right)$$

Proof continued

- Let $c_1 = P' - 2(m-1)$ and $c_2 = d_{m+1}$.
- Since $\lambda > \sqrt{k}$ and $c_1, c_2 \geq 2$,

$$(\lambda^{c_1-1} - \lambda)(\lambda^{c_2-1} - \lambda) \geq 0,$$

$$\lambda^{c_1+c_2-2} + \lambda^2 \geq \lambda^{c_1} + \lambda^{c_2},$$

$$-k(\lambda^{c_1} + \lambda^{c_2}) \geq -k(\lambda^{c_1+c_2-2} + \lambda^2),$$

$$(\lambda^{c_1} - k)(\lambda^{c_2} - k) \geq (\lambda^{c_1+c_2-2} - k)(\lambda^2 - k),$$

$$\frac{1}{(\lambda^{c_1} - k)(\lambda^{c_2} - k)} \leq \frac{1}{(\lambda^{c_1+c_2-2} - k)(\lambda^2 - k)}$$

Proof continued

- Substituting the c_i's,

$$\frac{1}{(\lambda^{P'-2(m-1)}-k)(\lambda^{d_{m+1}}-k)} \leq \frac{1}{(\lambda^{P'-2m+d_{m+1}}-k)(\lambda^2-k)}$$

- Multiplying the inductive hypothesis by $\frac{1}{\lambda^{d_{m+1}}-k}$,

$$\prod_{i=1}^{m+1} \frac{1}{\lambda^{d_i}-k} \leq \left(\frac{1}{\lambda^2-k}\right)^{m-1} \left(\frac{1}{\lambda^{P'-2(m-1)}-k}\right) \frac{1}{\lambda^{d_{m+1}}-k}$$

- Substituting the above inequality,

$$\prod_{i=1}^{m+1} \frac{1}{\lambda^{d_i}-k} \leq \left(\frac{1}{\lambda^2-k}\right)^{m} \left(\frac{1}{\lambda^{P'+d_{m+1}-2m}-k}\right)$$

$$= \left(\frac{1}{\lambda^2-k}\right)^{(m+1)-1} \left(\frac{1}{\lambda^{P-2((m+1)-1)}-k}\right)$$

The remaining arguments are based on those of Rampersad's, but add additional analysis to obtain the optimal bounds.

Lemma

Let m be an integer and p be a pattern over $\Delta = \{A_1, \ldots, A_m\}$. Suppose that for $1 \le i \le m$, A_i occurs $d_i \ge 2$ times in p.

1. *If $m \ge 3$ and $|p| \ge 4m$, then for $n \ge 0$, there are at least $(1.92)^n$ words of length n over 2 letters that avoid p.*

2. *If $m \ge 2$ and $|p| \ge 12$, then for $n \ge 0$, there are at least $(2.92)^n$ words of length n over 3 letters that avoid p.*

Proof

- Define S to be the set of all words over an alphabet Σ of size $k \in \{2,3\}$ that are instances of the pattern p.
- By Bell and Goh's lemma, the number of words of length n in S is at most $[x^n]C(x)$, where

$$C(x) := \sum_{i_1 \geq 1} \cdots \sum_{i_m \geq 1} k^{i_1+\cdots+i_m} x^{d_1 i_1+\cdots+d_m i_m}$$

- Define $B(x) := \sum_{i \geq 0} b_i x^i = (1 - kx + C(x))^{-1}$
  Set $\lambda = k - 0.08$. Clearly $b_0 = 1$ and $b_1 = k$. We show that $b_n \geq \lambda b_{n-1}$ for all $n \geq 1$, hence $b_n \geq \lambda^n$ for all $n \geq 0$.
- Then all coefficients of B are non-negative, thus Rampersad's theorem implies there are at least $b_n \geq \lambda^n$ words of length n having no factors in S, thus avoiding p.

Proof continued ($b_n \geq \lambda b_{n-1}$ for all $n \geq 1$)

- By induction on n, suppose $b_j \geq \lambda b_{j-1}$ for all $1 \leq j < n$.
- Expanding the left hand side of $B(x)(1 - kx + C(x)) = 1$,

$$\left(\sum_{i\geq 0} b_i x^i\right)\left(1 - kx + \sum_{i_1\geq 1}\cdots\sum_{i_m\geq 1} k^{i_1+\cdots+i_m} x^{d_1 i_1+\cdots+d_m i_m}\right)$$

- Hence for $n \geq 1$, $[x^n]B(x)(1 - kx + C(x)) = 0$, i.e.,

$$b_n - kb_{n-1} + \sum_{i_1\geq 1}\cdots\sum_{i_m\geq 1} k^{i_1+\cdots+i_m} b_{n-(d_1 i_1+\cdots+d_m i_m)} = 0$$

- Complete the induction by showing the major equation

$(k - \lambda)b_{n-1} - \sum_{i_1\geq 1}\cdots\sum_{i_m\geq 1} k^{i_1+\cdots+i_m} b_{n-(d_1 i_1+\cdots+d_m i_m)} \geq 0$

Proof continued

- Because $b_j \geq \lambda b_{j-1}$ for $1 \leq j < n$, $b_{n-i} \leq b_{n-1}/\lambda^{i-1}$ for $1 \leq i \leq n$. Therefore,

$$\sum_{i_1 \geq 1} \cdots \sum_{i_m \geq 1} k^{i_1 + \cdots + i_m} b_{n-(d_1 i_1 + \cdots + d_m i_m)}$$

$$\leq \lambda b_{n-1} \sum_{i_1 \geq 1} \frac{k^{i_1}}{\lambda^{d_1 i_1}} \cdots \sum_{i_m \geq 1} \frac{k^{i_m}}{\lambda^{d_m i_m}}$$

- Since $d_j \geq 2$ for $1 \leq j \leq m$, $k \leq 3$, and $\lambda > \sqrt{3}$,

$$\frac{k}{\lambda^{d_j}} \leq \frac{3}{\lambda^2} < 1$$

  thus all the geometric series converge.
- Computing the result, for $1 \leq j \leq m$,

$$\sum_{i_j \geq 1} \frac{k^{i_j}}{\lambda^{d_j i_j}} = \frac{k/\lambda^{d_j}}{1 - k/\lambda^{d_j}} = \frac{k}{\lambda^{d_j} - k}$$

Proof continued

- Thus

$$\sum_{i_1 \geq 1} \cdots \sum_{i_m \geq 1} k^{i_1 + \cdots + i_m} b_{n-(d_1 i_1 + \cdots + d_m i_m)} \leq k^m \lambda b_{n-1} \prod_{i=1}^{m} \frac{1}{\lambda^{d_i} - k}$$

- Applying our previous lemma to $P = |p|$, the key step is

$$\sum_{i_1 \geq 1} \cdots \sum_{i_m \geq 1} k^{i_1 + \cdots + i_m} b_{n-(d_1 i_1 + \cdots + d_m i_m)}$$
$$\leq k^m \lambda b_{n-1} \left(\frac{1}{\lambda^2 - k} \right)^{m-1} \left(\frac{1}{\lambda^{|p|-2(m-1)} - k} \right)$$

- It thus suffices to show the final inequality

$$(k - \lambda) \geq \lambda k^m \left(\frac{1}{\lambda^2 - k} \right)^{m-1} \left(\frac{1}{\lambda^{|p|-2(m-1)} - k} \right)$$

since multiplying this by b_{n-1} and using the key step derives the major equation.

Proof continued (Statement 1)

- The right hand side of the final inequality decreases as $|p|$ increases, thus it suffices to verify the case $|p| = 4m$. The final inequality is easily verified for $m = 3$ and $|p| = 12$.
- Now consider an arbitrary $m' \geq 3$ and p' with $|p'| = 4m'$. Substituting $\lambda = 1.92$ and $k = 2$, it follows that

$$c := \left(\frac{k}{\lambda^2 - k}\right)^{m'-m} \left(\frac{\lambda^{|p|-2(m-1)} - k}{\lambda^{|p'|-2(m'-1)} - k}\right)$$

$$\leq (1.19)^{m'-m} \left(\frac{1}{\lambda^{2(m'-m)}}\right) < 1$$

- Thus we conclude

$$k - \lambda \geq c\lambda k^m \left(\frac{1}{\lambda^2 - k}\right)^{m-1} \left(\frac{1}{\lambda^{|p|-2(m-1)} - k}\right)$$

$$= \lambda k^{m'} \left(\frac{1}{\lambda^2 - k}\right)^{m'-1} \left(\frac{1}{\lambda^{|p'|-2(m'-1)} - k}\right)$$

Proof continued (Statement 2)

For $m \geq 2$, it suffices to verify the final inequality for $|p| = \max\{12, 2m\}$.

- For $m = 2$ through $m = 5$ and $|p| = 12$, the equation is easily verified.
- For $m \geq 6$, $|p| = 2m$ and

$$\begin{aligned}
\lambda k^m \left(\frac{1}{\lambda^2 - k}\right)^{m-1} \left(\frac{1}{\lambda^{|p|-2(m-1)} - k}\right) &= 2.92 \left(\frac{3}{(2.92)^2 - 3}\right)^m \\
&\leq 2.92(0.5429)^m \\
&\leq 2.92(0.5429)^6 \\
&= 0.07476\cdots \\
&< 0.08 = k - \lambda
\end{aligned}$$

$\square$

Main results (strict bounds)

Both bounds below are strict in the sense that for every positive integer m, there exists a 2-unavoidable pattern with m distinct variables and length $3(2^{m-1}) - 1$ as well as a 3-unavoidable pattern with m distinct variables and length $2^m - 1$.

Theorem

Let p be a pattern with m distinct variables.

1. *If $|p| \geq 3(2^{m-1})$, then p is 2-avoidable.*
2. *If $|p| \geq 2^m$, then p is 3-avoidable.*

Proof (Statement 1)

We show by induction on m that if p is 2-unavoidable, $|p| < 3(2^{m-1})$.

- For $m = 1$, note that A^3 is 2-avoidable, hence A^ℓ is 2-avoidable for all $\ell \geq 3$. Thus if a unary pattern p is 2-unavoidable, $|p| < 3 = 3(2^{1-1})$.
- For $m = 2$, it is known that all binary patterns of length 6 are 2-avoidable (Roth 1992), hence all binary patterns of length at least 6 are also 2-avoidable. Thus if a binary pattern p is 2-unavoidable, $|p| < 6 = 3(2^{2-1})$.
- Now assume the statement holds for $m \geq 2$ and suppose p is a 2-unavoidable pattern with $m + 1$ distinct variables. For the sake of contradiction, assume that $|p| \geq 3(2^m)$.

Proof continued (Statement 1)

- Suppose p has a variable A that occurs exactly once. Let $p = p_1 A p_2$, where p_1 and p_2 are patterns with at most m variables. Without loss of generality, suppose $|p_1| \geq |p_2|$. Since $|p| \geq 3(2^m)$,

$$|p_1| \geq \left\lceil \frac{|p| - 1}{2} \right\rceil \geq \left\lceil \frac{3(2^m) - 1}{2} \right\rceil = 3(2^{m-1})$$

  By the contrapositive of the inductive hypothesis, p_1 is 2-avoidable. But p_1 divides p, hence p is 2-avoidable, a contradiction.

- Suppose every variable in p occurs at least twice. Since $|p| \geq 3(2^m) \geq 4(m+1)$ for $m \geq 2$, the previous lemma indicates there are infinitely many words over 2 letters that avoid p, thus p is 2-avoidable, a contradiction.

□

5. Extension to partial words

- We apply the power series approach to obtain similar bounds for avoidability in partial words, sequences that may contain some unknown characters or holes, denoted by $\diamond$'s, which are compatible or match any letter in the alphabet.

$$\begin{array}{c} a \; \diamond \; b \; \diamond \; a \\ \uparrow \\ \diamond \; \diamond \; b \; a \; a \end{array} \qquad\qquad \begin{array}{c} a \; \diamond \; b \; \diamond \; a \\ \not\uparrow \\ \diamond \; \diamond \; a \; a \; a \end{array}$$

- The modifications include that now we must avoid all partial words compatible with instances of the pattern. Lots of additional work with inequalities is necessary.

Partial word avoidability

- A partial word w over Σ is an instance of a pattern p over Δ if there exists a non-erasing morphism $\varphi : \Delta^* \to \Sigma^*$ such that $\varphi(p) \uparrow w$; the partial word w avoids p if none of its factors is an instance of p.

  $\underline{aa}\ \underline{b}\ \underline{a\diamond}\ c$ contains an instance of ABA while it avoids AAA
- A pattern p is called k-avoidable in partial words if for every $h \in \mathbb{N}$ there is a partial word with h holes over k letters avoiding p, or, equivalently, if there is a partial word over k letters with infinitely many holes which avoids p.
- The avoidability index for partial words is defined analogously to that of full words.

An upper bound

Lemma

Let $m \geq 1$ be an integer and p be a pattern over an alphabet $\Delta = \{A_1, \ldots, A_m\}$. Suppose that for $1 \leq i \leq m$, the variable A_i occurs $d_i \geq 1$ times in p. Let $k \geq 2$ be an integer and let Σ be a k-letter alphabet. Then for $n \geq 1$, the number of partial words of length n over Σ that are compatible with instances of the pattern p is no more than $[x^n]C(x)$, where

$$C(x) := \sum_{i_1 \geq 1} \cdots \sum_{i_m \geq 1} \left(\prod_{j=1}^{m} \left(k(2^{d_j} - 1) + 1 \right)^{i_j} \right) x^{d_1 i_1 + \cdots + d_m i_m}$$

A technical inequality

Lemma

Suppose $(k, \lambda) \in \{(2, 2.97), (3, 3.88)\}$ *and* $m \geq 1$ *is an integer. For any integer* P *and integers* d_j *for* $1 \leq j \leq m$ *such that* $d_j \geq 2$ *and* $P = d_1 + \cdots + d_m$,

$$\prod_{i=1}^{m} \frac{k(2^{d_i}-1)+1}{\lambda^{d_i}-(k(2^{d_i}-1)+1)} \leq \left(\frac{3k+1}{\lambda^2-(3k+1)}\right)^{m-1} \left(\frac{k}{(\frac{\lambda}{2})^{P-2(m-1)}-k}\right)$$

Exponential lower bounds

Lemma

Let $m \geq 4$ be an integer and p be a pattern over an alphabet $\Delta = \{A_1, \ldots, A_m\}$. Suppose that for $1 \leq i \leq m$, A_i occurs $d_i \geq 2$ times in p.

1. *If $|p| \geq 15(2^{m-3})$, then for $n \geq 0$, there are at least $(2.97)^n$ partial words of length n over 2 letters that avoid p.*
2. *If $|p| \geq 2^m$, then for $n \geq 0$, there are at least $(3.88)^n$ partial words of length n over 3 letters that avoid p.*

Arbitrarily many holes lemma

Thus for certain patterns, there exist λ^n partial words of length n that avoid the pattern, for some λ. It is not immediately clear that this is enough to prove the patterns are avoidable in partial words. The next lemma asserts this count is so large that it must include partial words with arbitrarily many holes, thus the patterns are 2-avoidable or 3-avoidable in partial words.

Lemma

Suppose $k \geq 2$ is an integer, $k < \lambda < k + 1$, Σ is an alphabet of size k, and S is a set of partial words over Σ with at least λ^n words of length n for each $n > 0$. For all integers $h \geq 0$, S contains a partial word with at least h holes.

- Unfortunately, the pattern $A^2BA^2CA^2$ of length $8 = 2^3$ is unavoidable in partial words (since some $a\diamond$ must occur infinitely often), thus to obtain the 2^m bound for avoidability as in the full word case, we require information about quaternary patterns of length $16 = 2^4$.
- Fortunately, for certain patterns, constructions can be made from full words avoiding a pattern to partial words avoiding a pattern that provide upper bounds on avoidability indices.

Bounds for partial words

Theorem

Let p be a pattern with m distinct variables.

1. *If* $m \geq 3$ *and* $|p| \geq 15(2^{m-3})$*, then p is 2-avoidable in partial words.*
2. *If* $m \geq 3$ *and* $|p| \geq 5(2^{m-2})$*, then p is 3-avoidable in partial words.*
3. *If* $m \geq 4$ *and* $|p| \geq 2^m$*, then p is 4-avoidable in partial words.*

3 gives a strict bound for 4-avoidability in partial words

Proof (Statement 3)

We show by induction on m that if p is 4-unavoidable, $|p| < 2^m$.

- We first establish the base case $m = 4$ by showing that every pattern p of length $16 = 2^4$ is 4-avoidable.
- Using the data in Blanchet-Sadri, Lohr and Scott 2012, the ternary patterns of length at least 7 which have avoidability index greater than 4 are

  $A^2BA^2CA^2$, of length 8

  A^2BA^2CA, A^2BACA^2, A^2BCA^2B, $\ldots$ of length 7

  (up to reversal and renaming of variables).

Blanchet-Sadri, F., Lohr, A., Scott, S.: Computing the partial word avoidability indices of ternary patterns. In Arumugam, S., Smyth, B., eds.: *IWOCA 2012, 23rd Int'l Workshop on Combinatorial Algorithms*. Vol. 7643 of *LNCS*, Berlin, Heidelberg, Springer-Verlag (2012) 206–218

Proof continued (Statement 3)

- If every variable in p occurs at least twice, our exponential lower bounds imply there exists a set S with at least $(3.88)^n$ ternary partial words of length n that avoid p for each $n \geq 0$. Applying our arbitrarily many holes lemma to S, for each $h \geq 0$, there exists a ternary partial word with at least h holes that avoids p. Thus p is 3-avoidable.
- Otherwise, p contains a variable α that occurs exactly once and $p = p_1 \alpha p_2$ for patterns p_1 and p_2 with at most 3 distinct variables. Note that $|p_1| + |p_2| = 15$.
- If p_1 has length at least 9, then p_1 is 4-avoidable, hence p is 4-avoidable by divisibility (likewise for p_2).
- Thus the only remaining case is when $|p_1| = 8$ and $|p_2| = 7$ (or vice versa).

Proof continued (Statement 3)

- If p_1 or p_2 is not in the list of ternary patterns mentioned before, it is 4-avoidable, hence p is 4-avoidable.
- Otherwise $p_1 = A^2BA^2CA^2$ up to a renaming of the variables. Note that p_1 contains a factor of the form A^2BA and all of the possible values of p_2 are on three variables, so they must contain B. This fits the form of a result of Blanchet-Sadri et al. which implies p is 4-avoidable.
- For $m \geq 5$, our exponential lower bounds and our arbitrarily many holes lemma imply that every pattern with length at least 2^m in which each variable appears at least twice is 3-avoidable.
- If p has a variable that occurs exactly once, we reason as in the proof of our main results to complete the induction.

$\square$

6. Conclusion

- Building upon the work of Rampersad 2011 and the power series techniques of Bell and Goh 2007, we have proved Cassaigne's 1994 conjecture that any pattern p with m distinct variables such that $|p| \geq 3(2^{m-1})$ is 2-avoidable, and any pattern p with m distinct variables such that $|p| \geq 2^m$ is 3-avoidable.
- Using in addition results and data about partial word avoidability of patterns from Blanchet-Sadri, Lohr and Scott 2012, we have also obtained exponential lower bounds for 2, 3 and 4-avoidability in partial words, the latter bound being strict.
- We do not know if our bounds for 2 and 3-avoidability in partial words are strict.

Thank you!