Random Strict Superpatterns

Martha Liendo, Anant Godbole

Department of Mathematics and Statistics
East Tennessee State University

June 12, 2012

Outline

- Introduction
- Cardinalities of Preferential Arrangements
- Containment Lengths of Strict Superpatterns
- Open Problems

A string, or word, contains a pattern if any *order-isomorphic* subsequence to that pattern can be found within that word. For example, the word 5371473 contains the subsequences 571, 574, and 473, each of which is order-isomorphic to the pattern 231, i.e. ordered the same as 231.

It also contains the subsequences 373 and 343 which are both order-isomorphic to the pattern 121. We call the patterns 121 and 231 unique *preferential arrangements* contained in the word.

Our interest lies in the word size required for a random word to contain all preferential arrangements of a given length. Such a string is known as a *superpattern*.

A string, or word, contains a pattern if any *order-isomorphic* subsequence to that pattern can be found within that word. For example, the word 5371473 contains the subsequences 571, 574, and 473, each of which is order-isomorphic to the pattern 231, i.e. ordered the same as 231.

It also contains the subsequences 373 and 343 which are both order-isomorphic to the pattern 121. We call the patterns 121 and 231 unique *preferential arrangements* contained in the word.

Our interest lies in the word size required for a random word to contain all preferential arrangements of a given length. Such a string is known as a *superpattern*.

A string, or word, contains a pattern if any *order-isomorphic* subsequence to that pattern can be found within that word. For example, the word 5371473 contains the subsequences 571, 574, and 473, each of which is order-isomorphic to the pattern 231, i.e. ordered the same as 231.

It also contains the subsequences 373 and 343 which are both order-isomorphic to the pattern 121. We call the patterns 121 and 231 unique *preferential arrangements* contained in the word.

Our interest lies in the word size required for a random word to contain all preferential arrangements of a given length. Such a string is known as a *superpattern*.

Some of the available work on placing bounds on the word length of superpatterns under permutation containment can be found in:

- M. H. Albert, M. D. Atkinson, C. C. Handley, D. A. Holton, W. Stromquist, On packing densities of permutations, 2002.
- H. Eriksson, K. Eriksson, S. Linusson, J. Wästlund, Dense packing of patterns in a permutation, 2002.
- A. Miller, Asymptotic bounds for permutations containing many different patterns, 2009.

Work by Burstein, Hästö, and Mansour (2003) studied the lengths of the shortest superpatterns under preferential arrangement containment. They offer a constructive proof for the minimum length required for a random word to be a strict superpattern.

Some of the available work on placing bounds on the word length of superpatterns under permutation containment can be found in:

- M. H. Albert, M. D. Atkinson, C. C. Handley, D. A. Holton, W. Stromquist, On packing densities of permutations, 2002.
- H. Eriksson, K. Eriksson, S. Linusson, J. Wästlund, Dense packing of patterns in a permutation, 2002.
- A. Miller, Asymptotic bounds for permutations containing many different patterns, 2009.

Work by Burstein, Hästö, and Mansour (2003) studied the lengths of the shortest superpatterns under preferential arrangement containment. They offer a constructive proof for the minimum length required for a random word to be a strict superpattern.

Let $[k] = \{1, 2, \ldots, k\}$ be a totally ordered alphabet of k letters, and let $[k]^n$ denote the set of all words of length n over this alphabet. For example, $[3]^2 = \{11, 22, 33, 12, 13, 21, 23, 32, 31\}$.

Order isomorphism partitions $[k]^n$ into a set of equivalence classes, where the equivalence class representative for each equivalence class, denoted here as π, is the word with the lowest possible ordinal numbers in the set of words which are all order-isomorphic to π and contained in $[k]^n$.

In the example above, the words 12, 13, and 23 are order-isomorphic and π is the word 12. The word π is called a *preferential arrangement*, or *ranking*. Of particular interest, is the set of all preferential arrangements of length n with $k \leq a$ ranks allowed, denoted as $\Pi(n, a)$.

Let $[k] = \{1, 2, \ldots, k\}$ be a totally ordered alphabet of k letters, and let $[k]^n$ denote the set of all words of length n over this alphabet. For example, $[3]^2 = \{11, 22, 33, 12, 13, 21, 23, 32, 31\}$.

Order isomorphism partitions $[k]^n$ into a set of equivalence classes, where the equivalence class representative for each equivalence class, denoted here as π, is the word with the lowest possible ordinal numbers in the set of words which are all order-isomorphic to π and contained in $[k]^n$.

In the example above, the words 12, 13, and 23 are order-isomorphic and π is the word 12. The word π is called a *preferential arrangement*, or *ranking*. Of particular interest, is the set of all preferential arrangements of length n with $k \leq a$ ranks allowed, denoted as $\Pi(n, a)$.

Let $[k] = \{1, 2, \ldots, k\}$ be a totally ordered alphabet of k letters, and let $[k]^n$ denote the set of all words of length n over this alphabet. For example, $[3]^2 = \{11, 22, 33, 12, 13, 21, 23, 32, 31\}$.

Order isomorphism partitions $[k]^n$ into a set of equivalence classes, where the equivalence class representative for each equivalence class, denoted here as π, is the word with the lowest possible ordinal numbers in the set of words which are all order-isomorphic to π and contained in $[k]^n$.

In the example above, the words 12, 13, and 23 are order-isomorphic and π is the word 12. The word π is called a *preferential arrangement*, or *ranking*. Of particular interest, is the set of all preferential arrangements of length n with $k \leq a$ ranks allowed, denoted as $\Pi(n, a)$.

Fact: For all values of $n \geq 1$ and $a \geq 1$, $\Pi(n,a) = \sum_{k=1}^{a} k!S(n,k)$, where $S(n,k) = \sum_{i=0}^{k}(-1)^i \binom{k}{i}(k-i)^n$ are Stirling numbers of the second kind.

Since $\Pi(n,a) = \Pi(a,a)$ when $n < a$, we are interested in cases where $n \geq a$, and most specifically where $n = a$.

In these cases $\Pi(a,a) = \sum_{k=1}^{a} k!S(a,k)$, the ordered Bell numbers.

Fact: For all values of $n \geq 1$ and $a \geq 1$,
$\Pi(n,a) = \sum_{k=1}^{a} k!S(n,k)$, where $S(n,k) = \sum_{i=0}^{k}(-1)^i\binom{k}{i}(k-i)^n$
are Stirling numbers of the second kind.

Since $\Pi(n,a) = \Pi(a,a)$ when $n < a$, we are interested in cases where $n \geq a$, and most specifically where $n = a$.

In these cases $\Pi(a,a) = \sum_{k=1}^{a} k!S(a,k)$, the ordered Bell numbers.

Fact: For all values of $n \geq 1$ and $a \geq 1$,
$\Pi(n,a) = \sum_{k=1}^{a} k! S(n,k)$, where $S(n,k) = \sum_{i=0}^{k} (-1)^i \binom{k}{i} (k-i)^n$
are Stirling numbers of the second kind.

Since $\Pi(n,a) = \Pi(a,a)$ when $n < a$, we are interested in cases where $n \geq a$, and most specifically where $n = a$.

In these cases $\Pi(a,a) = \sum_{k=1}^{a} k! S(a,k)$, the ordered Bell numbers.

Further characterization of superpatterns is necessary for clarity. Let a *minimal superpattern* be a superpattern in which no two adjacent letters are the same. Then a *minimum superpattern* is a minimal superpattern of the shortest length possible in which every letter is necessary for the containment of all preferential arrangements.

Let a *strict superpattern* be a superpattern in which the last letter of the superpattern is needed to complete one of the preferential arrangements contained in the superpattern. Clearly, all minimum superpatterns are strict superpatterns.

Following the notation of Burstein et al., let $n(a,a)$ be the length of a strict superpattern for $[a]^a$.

Further characterization of superpatterns is necessary for clarity. Let a *minimal superpattern* be a superpattern in which no two adjacent letters are the same. Then a *minimum superpattern* is a minimal superpattern of the shortest length possible in which every letter is necessary for the containment of all preferential arrangements.

Let a *strict superpattern* be a superpattern in which the last letter of the superpattern is needed to complete one of the preferential arrangements contained in the superpattern. Clearly, all minimum superpatterns are strict superpatterns.

Following the notation of Burstein et al., let $n(a,a)$ be the length of a strict superpattern for $[a]^a$.

Further characterization of superpatterns is necessary for clarity. Let a *minimal superpattern* be a superpattern in which no two adjacent letters are the same. Then a *minimum superpattern* is a minimal superpattern of the shortest length possible in which every letter is necessary for the containment of all preferential arrangements.

Let a *strict superpattern* be a superpattern in which the last letter of the superpattern is needed to complete one of the preferential arrangements contained in the superpattern. Clearly, all minimum superpatterns are strict superpatterns.

Following the notation of Burstein et al., let $n(a,a)$ be the length of a strict superpattern for $[a]^a$.

We sought to discover the wait time, W, necessary for a random word to be a strict superpattern. The binary case is trivial.

$$p_2(n) = p(W = n) = \frac{n-2}{2^{n-1}}$$

$$E(W) = 4$$

$$G_2(x) = \frac{x^3}{(2-x)^2}.$$

The case for $n(4,4)$ is rather problematic. Thus, the $n(3,3)$ case is studied.

We sought to discover the wait time, W, necessary for a random word to be a strict superpattern. The binary case is trivial.

$$p_2(n) = p(W = n) = \frac{n-2}{2^{n-1}}$$

$$E(W) = 4$$

$$G_2(x) = \frac{x^3}{(2-x)^2}.$$

The case for $n(4,4)$ is rather problematic. Thus, the $n(3,3)$ case is studied.

A superpattern for $[3]^3$ is a word that contains all 13 preferential arrangements of $[3]^3$, namely 111, 112, 121, 211, 122, 212, 221, 123, 132, 213, 231, 312, and 321.

Lemma *Any superpattern for* $[3]^3$ *contains a* jk *and a* kj *pattern both before and after at least one* i*, where* $i, j, k \in [3]$ *with* $i \neq j \neq k$.

Proof. Let σ be a superpattern for $[3]^3$ and let $i, j, k \in [3]$ with $i \neq j \neq k$.
Case 1: Assume σ does not contain a jk pattern before an i. Then σ does not contain the pattern jki and σ is not a superpattern for $[3]^3$. This is a contradiction and therefore σ contains a jk pattern before at least one i.
The cases for σ containing a jk pattern after an i, kj pattern before an i, and kj pattern after an i follow in a similar manner. □

A superpattern for $[3]^3$ is a word that contains all 13 preferential arrangements of $[3]^3$, namely 111, 112, 121, 211, 122, 212, 221, 123, 132, 213, 231, 312, and 321.

Lemma *Any superpattern for $[3]^3$ contains a jk and a kj pattern both before and after at least one i, where $i, j, k \in [3]$ with $i \neq j \neq k$.*

Proof. Let σ be a superpattern for $[3]^3$ and let $i, j, k \in [3]$ with $i \neq j \neq k$.
Case 1: Assume σ does not contain a jk pattern before an i. Then σ does not contain the pattern jki and σ is not a superpattern for $[3]^3$. This is a contradiction and therefore σ contains a jk pattern before at least one i.
The cases for σ containing a jk pattern after an i, kj pattern before an i, and kj pattern after an i follow in a similar manner. □

A superpattern for $[3]^3$ is a word that contains all 13 preferential arrangements of $[3]^3$, namely 111, 112, 121, 211, 122, 212, 221, 123, 132, 213, 231, 312, and 321.

Lemma *Any superpattern for* $[3]^3$ *contains a jk and a kj pattern both before and after at least one* i, *where* $i, j, k \in [3]$ *with* $i \neq j \neq k$.

Proof. Let σ be a superpattern for $[3]^3$ and let $i, j, k \in [3]$ with $i \neq j \neq k$.
Case 1: Assume σ does not contain a jk pattern before an i. Then σ does not contain the pattern jki and σ is not a superpattern for $[3]^3$. This is a contradiction and therefore σ contains a jk pattern before at least one i.
The cases for σ containing a jk pattern after an i, kj pattern before an i, and kj pattern after an i follow in a similar manner. □

The minimum length for $n(3,3)$ is given by Burstein et al. as $n = n(3,3) = 7$.

Theorem *There exist seven strict minimal superpatterns of length $n = 7$ up to isomorphism.*

Proof. The integer 7 can be partitioned into 3 parts in four ways, namely $(5,1,1)$, $(4,2,1)$, $(3,3,1)$, and $(3,2,2)$.

Case 1: Consider a strict minimal superpattern $\sigma = \sigma(1), \sigma(2), \ldots, \sigma(7) \in (5,1,1)$. Then there exists an $a_i > \lceil \frac{n}{2} \rceil = 4$, causing two adjacent letters to be the same letter, which contradicts σ as a strict minimal superpattern. Therefore there is no strict minimal superpattern $\sigma \in (5,1,1)$.

The minimum length for $n(3,3)$ is given by Burstein et al. as $n = n(3,3) = 7$.

Theorem *There exist seven strict minimal superpatterns of length $n = 7$ up to isomorphism.*

Proof. The integer 7 can be partitioned into 3 parts in four ways, namely $(5,1,1)$, $(4,2,1)$, $(3,3,1)$, and $(3,2,2)$.

Case 1: Consider a strict minimal superpattern $\sigma = \sigma(1), \sigma(2), \ldots, \sigma(7) \in (5,1,1)$. Then there exists an $a_i > \lceil \frac{n}{2} \rceil = 4$, causing two adjacent letters to be the same letter, which contradicts σ as a strict minimal superpattern. Therefore there is no strict minimal superpattern $\sigma \in (5,1,1)$.

The minimum length for $n(3,3)$ is given by Burstein et al. as $n = n(3,3) = 7$.

Theorem *There exist seven strict minimal superpatterns of length $n = 7$ up to isomorphism.*

Proof. The integer 7 can be partitioned into 3 parts in four ways, namely $(5,1,1)$, $(4,2,1)$, $(3,3,1)$, and $(3,2,2)$.

Case 1: Consider a strict minimal superpattern $\sigma = \sigma(1), \sigma(2), \ldots, \sigma(7) \in (5,1,1)$. Then there exists an $a_i > \lceil \frac{n}{2} \rceil = 4$, causing two adjacent letters to be the same letter, which contradicts σ as a strict minimal superpattern. Therefore there is no strict minimal superpattern $\sigma \in (5,1,1)$.

The minimum length for $n(3,3)$ is given by Burstein et al. as $n = n(3,3) = 7$.

Theorem *There exist seven strict minimal superpatterns of length $n = 7$ up to isomorphism.*

Proof. The integer 7 can be partitioned into 3 parts in four ways, namely $(5,1,1)$, $(4,2,1)$, $(3,3,1)$, and $(3,2,2)$.

Case 1: Consider a strict minimal superpattern $\sigma = \sigma(1), \sigma(2), \ldots, \sigma(7) \in (5,1,1)$. Then there exists an $a_i > \lceil \frac{n}{2} \rceil = 4$, causing two adjacent letters to be the same letter, which contradicts σ as a strict minimal superpattern. Therefore there is no strict minimal superpattern $\sigma \in (5,1,1)$.

Case 2: Consider a strict minimal superpattern $\sigma = \sigma(1), \sigma(2), \ldots, \sigma(7) \in (4,2,1)$ with $a_i = 1$, $a_j = 4$, and $a_k = 2$. Let $\sigma(a) = i$ for some $a \in [7]$. Then $a \geq 4$ since there exists both a jk and a kj pattern before i and $a \leq 4$ since there exists both a jk and a kj pattern after i. Therefore $a = 4$. Without loss of generality let $\sigma(1) = \sigma(3) = j$ and $\sigma(2) = k$. Since no two adjacent letters are the same letter, $\sigma(5) = \sigma(7) = j$ and $\sigma(6) = k$. Therefore there exists one such strict minimal superpattern $\sigma \in (4,2,1)$ up to isomorphism.

The cases for a strict minimal superpattern $\sigma \in (3,3,1)$ and a strict minimal superpattern $\sigma \in (3,2,2)$ follow in a similar manner. □

Case 2: Consider a strict minimal superpattern $\sigma = \sigma(1), \sigma(2), \ldots, \sigma(7) \in (4, 2, 1)$ with $a_i = 1$, $a_j = 4$, and $a_k = 2$. Let $\sigma(a) = i$ for some $a \in [7]$. Then $a \geq 4$ since there exists both a jk and a kj pattern before i and $a \leq 4$ since there exists both a jk and a kj pattern after i. Therefore $a = 4$. Without loss of generality let $\sigma(1) = \sigma(3) = j$ and $\sigma(2) = k$. Since no two adjacent letters are the same letter, $\sigma(5) = \sigma(7) = j$ and $\sigma(6) = k$. Therefore there exists one such strict minimal superpattern $\sigma \in (4, 2, 1)$ up to isomorphism.

The cases for a strict minimal superpattern $\sigma \in (3, 3, 1)$ and a strict minimal superpattern $\sigma \in (3, 2, 2)$ follow in a similar manner. □

The seven unique strict minimal superpatterns up to isomorphism of length $n = 7$ are 1213121, 1213212, 1231213, 1231231, 1231321, 1232123, and 1232132. Since the alphabet size is 3, there are 3! ways to permute the letters isomorphically in each strict minimal superpattern of length $n = 7$, giving a total of $3!(7) = 42$ strict minimal superpatterns of length $n = 7$.

Note: Burstein et al. show $n(\ell,\ell) \leq \ell^2 - 2\ell + 4$ and conjecture that $n(\ell,\ell) = \ell^2 - 2\ell + 4$. The previous theorem, along with lemmas proving no strict minimal superpattern exist on lengths $n < 7$ supports this for the case of $n(3,3)$.

The seven unique strict minimal superpatterns up to isomorphism of length $n = 7$ are 1213121, 1213212, 1231213, 1231231, 1231321, 1232123, and 1232132. Since the alphabet size is 3, there are 3! ways to permute the letters isomorphically in each strict minimal superpattern of length $n = 7$, giving a total of $3!(7) = 42$ strict minimal superpatterns of length $n = 7$.

Note: Burstein et al. show $n(\ell, \ell) \leq \ell^2 - 2\ell + 4$ and conjecture that $n(\ell, \ell) = \ell^2 - 2\ell + 4$. The previous theorem, along with lemmas proving no strict minimal superpattern exist on lengths $n < 7$ supports this for the case of $n(3, 3)$.

For expected containment length, consider the total amount of possible minimal superpatterns up to isomorphism of any length n. Since all minimal superpatterns are comprised of an alternating pattern, then, up to isomorphism, the first two letters can be fixed as i and j for $i, j \in [3]$ with $i \neq j$. There exist 2^{n-2} total words on the remaining $n-2$ positions that have alternating patterns since each letter can be chosen from an alphabet size of 2. However, not all of these 2^{n-2} words will result in a superpattern of $[3]^3$ on length n.

Our method of approach is to count the words which do not create superpatterns and subtract that amount from the total word count. The following lemma proves essential in determining which of these words are superpatterns.

For expected containment length, consider the total amount of possible minimal superpatterns up to isomorphism of any length n. Since all minimal superpatterns are comprised of an alternating pattern, then, up to isomorphism, the first two letters can be fixed as i and j for $i, j \in [3]$ with $i \neq j$. There exist 2^{n-2} total words on the remaining $n-2$ positions that have alternating patterns since each letter can be chosen from an alphabet size of 2. However, not all of these 2^{n-2} words will result in a superpattern of $[3]^3$ on length n.

Our method of approach is to count the words which do not create superpatterns and subtract that amount from the total word count. The following lemma proves essential in determining which of these words are superpatterns.

Lemma *Any strict minimal superpattern, σ, for $[3]^3$ of length $n \geq 7$ contains a minimum superpattern for $[3]^3$ with the last letter of the minimum superpattern occurring on the last letter of σ.*

Proof. Consider a strict minimal superpattern $\sigma = \sigma(1), \sigma(2), \ldots, \sigma(n)$ up to isomorphism for $[3]^3$ of length $n \geq 7$. Let $i, j, k \in [3]$ with $i < j < k$. Without loss of generality, let $\sigma(n) = i$ and $\sigma(n-1) = k$ since there are no two adjacent letters the same. Then there exists some $\sigma(b_1) = i$ as the first occurrence of i in σ, and, without loss of generality, there exists $\sigma(c_1), \sigma(c_2) = kj$ with $\sigma(c_2) = j$ as the last occurrence of j in σ where $b_1 < c_1 < c_2 < n-1$ since there exists both a jk and a kj pattern after at least one i and, up to isomorphism, it can be assumed that a kjk pattern satisfies this condition.

Lemma *Any strict minimal superpattern, σ, for $[3]^3$ of length $n \geq 7$ contains a minimum superpattern for $[3]^3$ with the last letter of the minimum superpattern occurring on the last letter of σ.*

Proof. Consider a strict minimal superpattern $\sigma = \sigma(1), \sigma(2), \ldots, \sigma(n)$ up to isomorphism for $[3]^3$ of length $n \geq 7$. Let $i, j, k \in [3]$ with $i < j < k$. Without loss of generality, let $\sigma(n) = i$ and $\sigma(n-1) = k$ since there are no two adjacent letters the same. Then there exists some $\sigma(b_1) = i$ as the first occurrence of i in σ, and, without loss of generality, there exists $\sigma(c_1), \sigma(c_2) = kj$ with $\sigma(c_2) = j$ as the last occurrence of j in σ where $b_1 < c_1 < c_2 < n-1$ since there exists both a jk and a kj pattern after at least one i and, up to isomorphism, it can be assumed that a kjk pattern satisfies this condition.

If $b_1 > 3$ then there exists a *jk* and a *kj* pattern before it, causing σ to contain either a *jkjikjk* or a *kjkikjk* pattern, both of which are strict superpatterns of length $n = 7$ and therefore $\sigma(n) = i$ is unnecessary for the the containment of all preferential arrangements. This contradicts the given fact that σ is a strict minimal superpattern. Therefore $b_1 \leq 3$.

Case 1: If $b_1 = 3$, then $\sigma(1), \sigma(2) = jk$ or kj. If $\sigma(1), \sigma(2) = jk$, then σ contains the minimum superpattern *jkikjki* with the last letter of the minimum superpattern occurring on the last letter of σ. If $\sigma(1), \sigma(2) = kj$, then σ contains the minimum superpattern *kjikjki* with the last letter of the minimum superpattern occurring on the last letter of σ.

The cases for $b_1 = 2$ and $b_1 = 1$ follow in a similar manner. □

If $b_1 > 3$ then there exists a jk and a kj pattern before it, causing σ to contain either a $jkjikjk$ or a $kjkikjk$ pattern, both of which are strict superpatterns of length $n = 7$ and therefore $\sigma(n) = i$ is unnecessary for the the containment of all preferential arrangements. This contradicts the given fact that σ is a strict minimal superpattern. Therefore $b_1 \leq 3$.

Case 1: If $b_1 = 3$, then $\sigma(1), \sigma(2) = jk$ or kj. If $\sigma(1), \sigma(2) = jk$, then σ contains the minimum superpattern $jkikjki$ with the last letter of the minimum superpattern occurring on the last letter of σ. If $\sigma(1), \sigma(2) = kj$, then σ contains the minimum superpattern $kjikjki$ with the last letter of the minimum superpattern occurring on the last letter of σ.

The cases for $b_1 = 2$ and $b_1 = 1$ follow in a similar manner. □

If $b_1 > 3$ then there exists a *jk* and a *kj* pattern before it, causing σ to contain either a *jkjikjk* or a *kjkikjk* pattern, both of which are strict superpatterns of length $n = 7$ and therefore $\sigma(n) = i$ is unnecessary for the the containment of all preferential arrangements. This contradicts the given fact that σ is a strict minimal superpattern. Therefore $b_1 \leq 3$.

Case 1: If $b_1 = 3$, then $\sigma(1), \sigma(2) = jk$ or kj. If $\sigma(1), \sigma(2) = jk$, then σ contains the minimum superpattern *jkikjki* with the last letter of the minimum superpattern occurring on the last letter of σ. If $\sigma(1), \sigma(2) = kj$, then σ contains the minimum superpattern *kjikjki* with the last letter of the minimum superpattern occurring on the last letter of σ.
The cases for $b_1 = 2$ and $b_1 = 1$ follow in a similar manner. □

Therefore, the words that fail to create a superpattern of $[3]^3$ do not contain a complete embedding of one of the strict minimal superpatterns of length seven.

All the words with alternating patterns contain some portion of a strict minimal superpattern of length seven up to isomorphism since the first two letters are fixed as 1 and 2 and each strict minimal superpattern of length seven can be written in the same manner. Let an i-fold progression count the amount of the 2^{n-2} words which contain the first through the ith letters of a unique strict minimal superpattern of length seven but not the $i+1$st letter.

Therefore, the words that fail to create a superpattern of $[3]^3$ do not contain a complete embedding of one of the strict minimal superpatterns of length seven.

All the words with alternating patterns contain some portion of a strict minimal superpattern of length seven up to isomorphism since the first two letters are fixed as 1 and 2 and each strict minimal superpattern of length seven can be written in the same manner. Let an i-fold progression count the amount of the 2^{n-2} words which contain the first through the ith letters of a unique strict minimal superpattern of length seven but not the $i+1$st letter.

Then 2-fold progression is guaranteed by the fixed 1 and 2 occurring on the first and second positions of each word. The third position must be an 1 or a 3 since no two adjacent letters are the same letter. Let the strict minimal superpatterns of length seven with the first three positions containing the pattern 121 be called type A and the remaining strict minimal superpatterns of length seven be called type B.

Type A:	1213121	Type B:	1231231
	1213212		1231321
			1231213
			1232123
			1232132

Then 2-fold progression is guaranteed by the fixed 1 and 2 occurring on the first and second positions of each word. The third position must be an 1 or a 3 since no two adjacent letters are the same letter. Let the strict minimal superpatterns of length seven with the first three positions containing the pattern 121 be called type A and the remaining strict minimal superpatterns of length seven be called type B.

Type A:	1213121	Type B:	1231231
	1213212		1231321
			1231213
			1232123
			1232132

First, consider the strict minimal superpatterns of type A. A word that satisfies 3-fold progression contains the pattern 121 on the first three positions, but no 3 afterwards. There is one such word, namely $121212\ldots$, which satisfies a 3-fold progression.

For a 4-fold progression to occur, the word must contain the pattern 121 on the first three positions followed by a 3 which has no 1 or 2 after it, otherwise a 5-fold progression will occur. There is only one such word, namely $121212\ldots3$, where the only occurrence of 3 is at the end of the word.

Since any other occurrence of 3 on the $(n-4)$ remaining positions results in a 5-fold progression, there are $n-4$ ways for the word to contain a 5-fold progression for each possible letter that can follow 3, giving $2(n-4)$ ways. These words follow the patterns $121212\ldots313131\ldots$ and $121212\ldots323232\ldots$.

First, consider the strict minimal superpatterns of type A. A word that satisfies 3-fold progression contains the pattern 121 on the first three positions, but no 3 afterwards. There is one such word, namely 121212 . . ., which satisfies a 3-fold progression.

For a 4-fold progression to occur, the word must contain the pattern 121 on the first three positions followed by a 3 which has no 1 or 2 after it, otherwise a 5-fold progression will occur. There is only one such word, namely 121212 . . . 3, where the only occurrence of 3 is at the end of the word.

Since any other occurrence of 3 on the $(n-4)$ remaining positions results in a 5-fold progression, there are $n-4$ ways for the word to contain a 5-fold progression for each possible letter that can follow 3, giving $2(n-4)$ ways. These words follow the patterns 121212 . . . 313131 . . . and 121212 . . . 323232 . . ..

First, consider the strict minimal superpatterns of type A. A word that satisfies 3-fold progression contains the pattern 121 on the first three positions, but no 3 afterwards. There is one such word, namely $121212\ldots$, which satisfies a 3-fold progression.

For a 4-fold progression to occur, the word must contain the pattern 121 on the first three positions followed by a 3 which has no 1 or 2 after it, otherwise a 5-fold progression will occur. There is only one such word, namely $121212\ldots3$, where the only occurrence of 3 is at the end of the word.

Since any other occurrence of 3 on the $(n-4)$ remaining positions results in a 5-fold progression, there are $n-4$ ways for the word to contain a 5-fold progression for each possible letter that can follow 3, giving $2(n-4)$ ways. These words follow the patterns $121212\ldots313131\ldots$ and $121212\ldots323232\ldots$.

In order for a word to contain a 6-fold progression, it must contain the 5-fold progression $121212\ldots313131\ldots$ followed by a 2 that is not followed by a 1 or $121212\ldots323232\ldots$ pattern followed by a 1 not followed by a 2. This corresponds to all the ways in which non-consecutive choices can be made on length $n-3$, so there are $2\binom{n-4}{2}$ such words, namely $121212\ldots313131\ldots232323\ldots$ and $121212\ldots323232\ldots131313\ldots$.

Therefore the total count for the amount of words which do not contain a complete embedding of one of the type A strict minimal superpatterns of length seven is

$$\beta_A(n) = 1 + 1 + 2(n-4) + 2\binom{n-4}{2} = n^2 - 7n + 14.$$

In order for a word to contain a 6-fold progression, it must contain the 5-fold progression $121212\ldots313131\ldots$ followed by a 2 that is not followed by a 1 or $121212\ldots323232\ldots$ pattern followed by a 1 not followed by a 2. This corresponds to all the ways in which non-consecutive choices can be made on length $n-3$, so there are $2\binom{n-4}{2}$ such words, namely $121212\ldots313131\ldots232323\ldots$ and $121212\ldots323232\ldots131313\ldots$.

Therefore the total count for the amount of words which do not contain a complete embedding of one of the type A strict minimal superpatterns of length seven is

$$\beta_A(n) = 1 + 1 + 2(n-4) + 2\binom{n-4}{2} = n^2 - 7n + 14.$$

The argument for the words failing to contain type B follows in a similar manner, giving

$$\beta_B(n) = 3n - 10.$$

Therefore the total amount of words that do not contain a complete embedding of one of the strict minimal superpatterns of length seven is

$$\beta_{total}(n) = (n-2)^2.$$

The argument for the words failing to contain type B follows in a similar manner, giving

$$\beta_B(n) = 3n - 10.$$

Therefore the total amount of words that do not contain a complete embedding of one of the strict minimal superpatterns of length seven is

$$\beta_{total}(n) = (n-2)^2.$$

The total amount of minimal superpatterns up to isomorphism of any length n

$$\Gamma_{total}(n) = 2^{n-2} - (n-2)^2.$$

The sequence generated by $\Gamma_{total}(n)$ exists as entry number A024012 in oeis, but with no context.

The total amount of minimal superpatterns up to isomorphism of any length n

$$\Gamma_{total}(n) = 2^{n-2} - (n-2)^2.$$

The sequence generated by $\Gamma_{total}(n)$ exists as entry number A024012 in oeis, but with no context.

Lemma *For all $n \geq 7$, the total amount of strict minimal superpattern of length n, $S_m(n) = (n-4)^2 - 2$.*

Proof. The amount of strict minimal superpattern of length n up to isomorphism will equal the total amount of minimal superpatterns up to isomorphism of length n minus any non-strict superpatterns of length n up to isomorphism. The total amount of non-strict superpatterns of length n up to isomorphism is equal to the total amount of minimal superpatterns up to isomorphism of length $n-1$ times 2, since the last letter is unnecessary in a non-strict superpattern for the completion of any preferential arrangement of $[3]^3$, making the word on the first $n-1$ letters a valid minimal superpattern of length $n-1$ and there are 2 choices for the nth letter since no two adjacent letters in the word are the same letter. Therefore,

Lemma *For all $n \geq 7$, the total amount of strict minimal superpattern of length n, $S_m(n) = (n-4)^2 - 2$.*

Proof. The amount of strict minimal superpattern of length n up to isomorphism will equal the total amount of minimal superpatterns up to isomorphism of length n minus any non-strict superpatterns of length n up to isomorphism. The total amount of non-strict superpatterns of length n up to isomorphism is equal to the total amount of minimal superpatterns up to isomorphism of length $n-1$ times 2, since the last letter is unnecessary in a non-strict superpattern for the completion of any preferential arrangement of $[3]^3$, making the word on the first $n-1$ letters a valid minimal superpattern of length $n-1$ and there are 2 choices for the nth letter since no two adjacent letters in the word are the same letter. Therefore,

$$\begin{aligned} S_m(n) &= [2^{n-2} - (n-2)^2] - 2[2^{n-3} - (n-3)^2] \\ &= 2^{n-2} - n^2 + 4n - 4 - 2^{n-2} + 2n^2 - 12n + 18 \\ &= n^2 - 8n + 14 \\ &= (n-4)^2 - 2. \end{aligned}$$

□

The sequence generated by $S_m(n)$ exists as entry number $A008865$ in oeis, but with very little context.

Now that we have counted the strict minimal superpatterns of length n we need to count the strict non-minimal superpatterns.

$$\begin{aligned} S_m(n) &= [2^{n-2} - (n-2)^2] - 2[2^{n-3} - (n-3)^2] \\ &= 2^{n-2} - n^2 + 4n - 4 - 2^{n-2} + 2n^2 - 12n + 18 \\ &= n^2 - 8n + 14 \\ &= (n-4)^2 - 2. \end{aligned}$$

□

The sequence generated by $S_m(n)$ exists as entry number $A008865$ in oeis, but with very little context.

Now that we have counted the strict minimal superpatterns of length n we need to count the strict non-minimal superpatterns.

$$\begin{aligned} S_m(n) &= [2^{n-2}-(n-2)^2]-2[2^{n-3}-(n-3)^2] \\ &= 2^{n-2}-n^2+4n-4-2^{n-2}+2n^2-12n+18 \\ &= n^2-8n+14 \\ &= (n-4)^2-2. \end{aligned}$$

□

The sequence generated by $S_m(n)$ exists as entry number $A008865$ in oeis, but with very little context.

Now that we have counted the strict minimal superpatterns of length n we need to count the strict non-minimal superpatterns.

Lemma *For all $n \geq 8$, the amount of strict non-minimal superpatterns of length n up to isomorphism,* $S_a(n) = \sum_{m=7}^{n-1}[(n-4)^2 - 2]\binom{n-2}{m-2}$.

Proof. A non-minimal superpattern can be reduced to a minimal superpattern, therefore any strict non-minimal superpattern of length n will contain an embedded occurrence of a strict minimal superpattern of length m, where $7 \leq m < n$. All such superpatterns are found by inserting $n-m$ letters which cause two adjacent letters to be the same into strict minimal superpatterns of length m.

Lemma *For all $n \geq 8$, the amount of strict non-minimal superpatterns of length n up to isomorphism,* $S_a(n) = \sum_{m=7}^{n-1}[(n-4)^2 - 2]\binom{n-2}{m-2}$.

Proof. A non-minimal superpattern can be reduced to a minimal superpattern, therefore any strict non-minimal superpattern of length n will contain an embedded occurrence of a strict minimal superpattern of length m, where $7 \leq m < n$. All such superpatterns are found by inserting $n - m$ letters which cause two adjacent letters to be the same into strict minimal superpatterns of length m.

These insertions can take place anywhere in the word except before the last letter since an occurrence of two adjacent letters as the same letter at the end of the word contradicts the fact that the last letter is necessary for the completion of at least one preferential arrangement of $[3]^3$ contained in that superpattern.

Therefore there are $n-m$ insertions into $n-1$ possible positions and there are $\binom{n-1}{n-m}$ ways to do this. Utilization of combination identities gives $\binom{n-1}{n-m} = \binom{n-1}{(n-1)-(m-1)} = \binom{n-1}{m-1}$.

These insertions can take place anywhere in the word except before the last letter since an occurrence of two adjacent letters as the same letter at the end of the word contradicts the fact that the last letter is necessary for the completion of at least one preferential arrangement of $[3]^3$ contained in that superpattern.

Therefore there are $n-m$ insertions into $n-1$ possible positions and there are $\binom{n-1}{n-m}$ ways to do this. Utilization of combination identities gives $\binom{n-1}{n-m} = \binom{n-1}{(n-1)-(m-1)} = \binom{n-1}{m-1}$.

The total amount ways to insert all $m-1$ letters into the $n-1$ positions is then $\binom{n-2}{m-2}$ and since this insertion method can be done for all strict minimal superpatterns of length m,

$$S_a(n) = \sum_{m=7}^{n-1} S_m(n)\binom{n-2}{m-2} = \sum_{m=7}^{n-1} [(n-4)^2-2]\binom{n-2}{m-2}.$$

□

The total amount ways to insert all $m-1$ letters into the $n-1$ positions is then $\binom{n-2}{m-2}$ and since this insertion method can be done for all strict minimal superpatterns of length m,

$$\begin{aligned} S_a(n) &= \sum_{m=7}^{n-1} S_m(n)\binom{n-2}{m-2} \\ &= \sum_{m=7}^{n-1} [(n-4)^2-2]\binom{n-2}{m-2}. \end{aligned}$$

□

Theorem *For all $n \geq 7$ the total amount of strict superpatterns of length n, $S(n) = 6\sum_{m=7}^{n}[(n-4)^2 - 2]\binom{n-2}{m-2}$.*

Thus, the wait time necessary for a random word on n letters to be a strict superpattern is

$$p_3(n) = p(W = n) = \frac{6}{3^n}\sum_{m=7}^{n}[(n-4)^2 - 2]\binom{n-2}{m-2}.$$

Theorem *For all $n \geq 7$ the total amount of strict superpatterns of length n, $S(n) = 6\sum_{m=7}^{n}[(n-4)^2 - 2]\binom{n-2}{m-2}$.*

Thus, the wait time necessary for a random word on n letters to be a strict superpattern is

$$p_3(n) = p(W = n) = \frac{6}{3^n}\sum_{m=7}^{n}[(n-4)^2 - 2]\binom{n-2}{m-2}.$$

By using different forms of the derivatives of the geometric series, the the expected wait time can be computed.

$$\begin{aligned} E(W) &= \sum_{n=7}^{\infty} \frac{6n}{3^n} \sum_{m=7}^{n} [(n-4)^2 - 2] \binom{n-2}{m-2} \\ &= 11.875. \end{aligned}$$

This validates a numerical analysis of the problem done in Matlab, which gives the expected wait time for the creation of a strict superpattern as 11.820 using 1000 trials.

By using different forms of the derivatives of the geometric series, the the expected wait time can be computed.

$$\begin{aligned} E(W) &= \sum_{n=7}^{\infty} \frac{6n}{3^n} \sum_{m=7}^{n} [(n-4)^2 - 2] \binom{n-2}{m-2} \\ &= 11.875. \end{aligned}$$

This validates a numerical analysis of the problem done in Matlab, which gives the expected wait time for the creation of a strict superpattern as 11.820 using 1000 trials.

The generating function is

$$\begin{aligned} G_3(x) &= \sum_{n=7}^{\infty} \frac{6t^n}{3^n} \sum_{m=7}^{n} [(n-4)^2 - 2] \binom{n-2}{m-2} \\ &= \frac{2t^7(16t^2 - 63t + 63)}{(3-t)^5(3-2t)^2}. \end{aligned}$$

This work for the cases of $n(2,2)$ and $n(3,3)$ creates open problems of interest. A few such question which require further investigation are:

(i) Is there a bijective proof for $\beta_{total}(n) = (n-2)^2$ versus the i-fold progression argument?

(ii) What are the lower bounds and containment length for cases of $n(4,4)$?

This work for the cases of $n(2,2)$ and $n(3,3)$ creates open problems of interest. A few such question which require further investigation are:

(i) Is there a bijective proof for $\beta_{total}(n) = (n-2)^2$ versus the i-fold progression argument?

(ii) What are the lower bounds and containment length for cases of $n(4,4)$?

This work for the cases of $n(2,2)$ and $n(3,3)$ creates open problems of interest. A few such question which require further investigation are:

(i) Is there a bijective proof for $\beta_{total}(n) = (n-2)^2$ versus the i-fold progression argument?

(ii) What are the lower bounds and containment length for cases of $n(4,4)$?

Some investigation into case (ii) has revealed that many difficulties will arise in the achievement of a counting method and the fold method of counting utilized in the $n(3, 3)$ case for containment will not work for the case of $n(4, 4)$.

One major complication to note is that a minimum superpattern for $[4]^4$ of length 12 can be constructed using the construction method found in work by Burstein et al., but there exist strict superpatterns for $[4]^4$ of lengths larger than 12 which do not contain one of the minimum superpatterns. One such example can be constructed following the construction of the type A strict superpatterns for $[3]^3$, i.e. 12131214121312 1.

Thank you.

Some investigation into case (ii) has revealed that many difficulties will arise in the achievement of a counting method and the fold method of counting utilized in the $n(3,3)$ case for containment will not work for the case of $n(4,4)$.

One major complication to note is that a minimum superpattern for $[4]^4$ of length 12 can be constructed using the construction method found in work by Burstein et al., but there exist strict superpatterns for $[4]^4$ of lengths larger than 12 which do not contain one of the minimum superpatterns. One such example can be constructed following the construction of the type A strict superpatterns for $[3]^3$, i.e. 12131214121312 1.

Thank you.

Some investigation into case (ii) has revealed that many difficulties will arise in the achievement of a counting method and the fold method of counting utilized in the $n(3,3)$ case for containment will not work for the case of $n(4,4)$.

One major complication to note is that a minimum superpattern for $[4]^4$ of length 12 can be constructed using the construction method found in work by Burstein et al., but there exist strict superpatterns for $[4]^4$ of lengths larger than 12 which do not contain one of the minimum superpatterns. One such example can be constructed following the construction of the type A strict superpatterns for $[3]^3$, i.e. 121312141213121.

Thank you.