Star Anagram Detection and Classification

Jason Parker[1] *and Dan Barker*[2]

[1] *Air Force Research Laboratory, Wright Patterson Air Force Base, USA,* `jason.parker.13@us.af.mil`
[2] *Freedom From Religion Foundation, Madison, WI, USA,* `dbarker@ffrf.org`

Abstract

A star anagram is a rearrangement of the letters of one word to produce another word where no letter retains its original neighbors. These maximally shuffled anagrams are rare, comprising only about 5.7% of anagrams in English. They can also be depicted as unicursal polygons with varying forms, including the eponymous stars. We develop automated methods for detecting stars among other anagrams and for classifying them based on their polygon's degree of both rotational and reflective symmetry. Next, we explore several properties of star anagrams including proofs for two results about the edge lengths of perfect, i.e., maximally symmetric, stars leveraging perhaps surprising connections to modular arithmetic and the celebrated Chinese Remainder Theorem. Finally, we conduct an exhaustive search of English for star anagrams and provide numerical results about their clustering into common shapes along with examples of geometrically noteworthy stars.

Keywords: Star anagram, Star polygon, Unicursal polygon, Symmetric polygon, Chinese Remainder Theorem.

1 Introduction

1.1 Motivation

An *anagram* is a word or phrase formed by rearranging the letters of another word or phrase. In this article, we use the word anagram to refer only to a pair of single words that are anagrams of each other. We consider the pair to be ordered and view the second word as a rearrangement of the letters in the first word, e.g., `EARTH` $\to$ `HEART`. Notice that `EARTH` $\to$ `HEART` is a simple rotation. All letters keep their original neighbors; nothing has been shuffled.

A *star anagram* is a special class of anagrams in which the letters have been maximally shuffled: no letter in the second word is adjacent to one of its original neighbors, counting the first and last letters as neighbors. An example is `EARTH` → `HATER`. The name *star* anagram derives from an interesting geometric property of these anagrams. In particular, if we arrange the letters of the first word in a circle and trace the path formed by the rearranged second word, we obtain a star as on the right in Figure 1. Most anagrams, like our earlier example `EARTH` → `HEART`, do not produce a star shape, as we see on the left of the Figure.

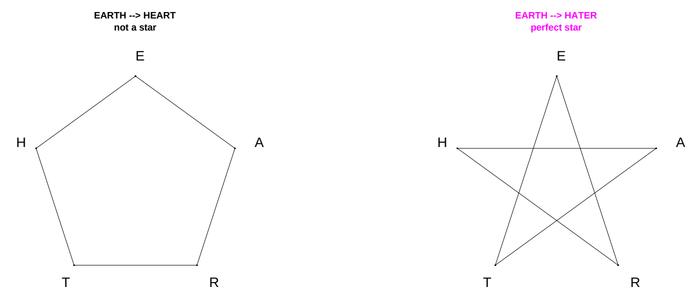

Figure 1: (Left) An anagram that is not a star. (Right) An example star anagram.

An interesting subset of star anagrams are *symmetric*. They can be folded perfectly along some dividing line (reflective symmetry) or rotated less than one full turn and look the same (rotational symmetry). Stars which lack this symmetry are *asymmetric*. An even rarer type of star anagram has all edges of the same length, which we denote as *perfect*. Note that perfect stars are both reflexively and rotationally symmetric, although we place them into a special class. Figure 2 provides examples of all three classes of star anagrams for words of length 8.

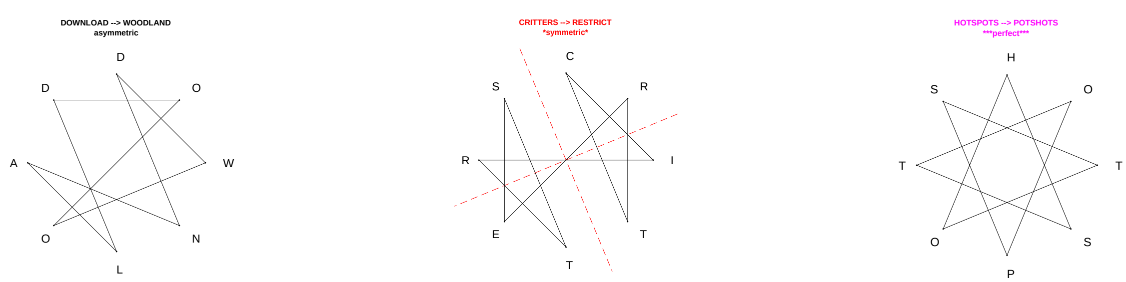

Figure 2: Length 8 examples of the three star anagrams classes.

Barker [Bar94] originally coined the term star anagram and introduced it to the first author. Prior to the work described in this article, Barker searched for star anagrams without automated tools. While intellectually rewarding (because it challenges you to turn one dimension into two), this approach makes it difficult to identify large groups of star anagrams, particularly among longer words and those with repeated letters.

1.2 Contribution

The primary contribution of this paper is a numerically inexpensive method for automatically detecting star anagrams and classifying them based on their degree of rotational and reflective symmetry, including a simple test for perfection. All of these methods rely on simple operations computed from the edge lengths of the anagram's representation as a unicursal polygon. We also use the Chinese Remainder Theorem to prove that perfect stars must have edge lengths that are coprime with their word length, a result already well known in the study of star polygons.

A star anagram's polygon changes based on the ordering of the words. We prove that reversing the order of the anagram preserves both starriness and perfection. A surprising result on the edge length of reversed perfect stars is also provided, demonstrating that the edge lengths of a perfect star and its reversed star are modular inverses in the parlance of number theory.

Finally, we conduct a detailed numerical study of the star anagrams in English. First, all star anagrams in a large database of English words are detected and classified. We then provide numerical results on the clustering of these star anagrams into common shapes and their distribution across word lengths. We also discuss the initially surprising notion of *autostars*, which are words that can be star anagrams of themselves. An exhaustive search of autostars provides interesting examples of polygon shapes that do not appear among normal star anagrams in the English language.

1.3 Outline

The remainder of this paper is organized as follows. Section 2 describes our approach for detecting star anagrams, and Section 3 describes our approach for star anagram classification. In Section 4, we prove several properties of star anagrams, discuss clustering of stars into common shapes, and introduce autostars. Section 5 presents the numerical results for our search of English words for star anagrams. Finally, Section 6 provides concluding remarks and possible future work.

1.4 Notation

Throughout this article we will use bold face capital letters for matrices (e.g., $\boldsymbol{A}$), bold face lower case letters for vectors (e.g., $\boldsymbol{p}$), and non-bold letters for scalars (e.g., N). We will denote the set of integers as $\mathbb{Z}$. A length N vector $\boldsymbol{p}$ of integers will be written as $\boldsymbol{p} \in \mathbb{Z}^N$, with the n^{th} entry denoted as p_n. Similarly, a matrix with M rows and N columns will be denoted as $\boldsymbol{A} \in \mathbb{Z}^{M \times N}$, with the scalar entry in the m^{th} row and n^{th} column denoted as a_{mn}. Note that we use 0 based indexing, e.g., numbering the columns from $0 \ldots N-1$, throughout this article.

We use $N!$ for the factorial of a scalar N, and the magnitude of a scalar p will be given as $|p|$. The modulo N operation (i.e., remainder after division by N) for a scalar integer K will be denoted as $K \bmod N$. We say that $a \equiv b \pmod{N}$ if $a \bmod N = b \bmod N$. Finally, we say that two integers N and L are *coprime* if they have no common positive divisor other than 1.

2 Star Anagram Detection

Generating a comprehensive list of all anagrams in a given set of words is straightforward, e.g., by exhaustively comparing the sorted letters of equal length words. Omitting those details, we will focus on an algorithmic approach for detecting whether a given anagram is a star. Before describing this approach, we need to explain how to think of anagrams as paths.

2.1 Anagrams as Paths

We number the letters of any length N word with the integer values 0 to $N-1$. Any rearrangement of these letters can be viewed as traversing a *path* that connects the letters in the specified order. We will represent such a path as a vector $\boldsymbol{p} \in \mathbb{Z}^N$ with entries $\{p_n\}_{n=0}^{N-1}$. For our example `EARTH` $\rightarrow$ `HATER`, we obtain the path $\boldsymbol{p} = [4, 1, 3, 0, 2]$. Note that for a given word of length N there are $N!$ possible paths, including the original word and many nonsense arrangements.

In our analysis of star anagrams, we think of the letters of the original word as nodes arranged uniformly around a circle or ring. The path is drawn as a series of line segments connecting the nodes in the specified order, including a segment from node p_{N-1} back to node p_0 to close the figure. These shapes produced by a continuous path are known in geometry as *unicursal polygons* and have been widely studied. Indeed, *star polygons*, which are the unicursal polygons produced by our star anagrams, have been studied since at least the fourteenth century [Cox69, Section 2.8]. In the sequel, we will see that star anagram detection and classification can be done entirely by looking at the properties of anagram paths.

2.2 Identifying All Possible Paths

Our simple example `EARTH` $\rightarrow$ `HATER` conveniently hid a complication. In particular, the path for an anagram with repeated letters is not unique. The nodes of the repeated letters can be swapped in the path without changing the resulting word. For an anagram with R repeated letters each of which appears w_r times, the number of possible paths P will be $P = \prod_{r=0}^{R-1} w_r!$, which can be large.

For example, consider `CAREERS` $\rightarrow$ `CREASER`. The two "e" and "r" letters can both be visited in either order in the path, leading to $P = (2! \cdot 2!) = 4$ possible paths as shown in Figure 3. Only the fourth path reveals this anagram to be a perfect star.

Our solution to this problem is straightforward. We simply generate all possible paths for each anagram using an exhaustive recursive enumeration and evaluate each path. We then select a single path $\boldsymbol{p}^*$ out of the set of P possible paths $\{\boldsymbol{p}_i\}_{i=0}^{P-1}$ to represent the anagram. A perfect star path is selected if found, followed in preference by a symmetric star path, an asymmetric star path, and finally a non-star path. If multiple paths are

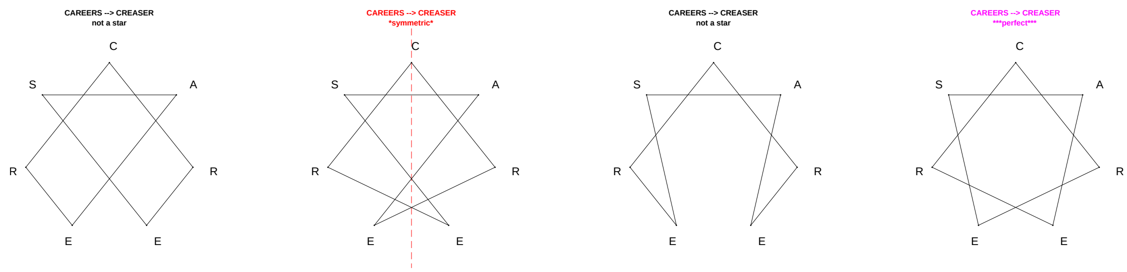

Figure 3: The 4 possible paths for the perfect star anagram `CAREERS` $\rightarrow$ `CREASER`.

found within the preferred class, we select one arbitrarily. Multiple symmetric star paths are handled differently, which we describe after discussing classification.

2.3 Step Sizes

Now that we have identified the set of paths to test for a given anagram, we turn our attention to computing the steps around the circle represented by a given path. First, given a path $\boldsymbol{p}$, we define the path differences $\boldsymbol{d} \in \mathbb{Z}^N$ with entries $\{d_n\}_{n=0}^{N-1}$ given by

$$d_n = p_{n+1} - p_n \tag{1}$$

where for convenience we define $p_N = p_0$. We would like to use these differences to analyze the geometric properties of the path around the circle.

However, these raw differences of the node locations include ambiguities that make direct analysis difficult. Notice that d_n can take on values from $-(N-1)$ to $N-1$, yielding $2N-2$ possible values[1]. Clearly these values are redundant, since starting from a given node there are only $N-1$ possible *steps* to the next node. Our goal will be to map these path differences to unambiguous steps.

If the next node is k_1 steps away around the circle in the clockwise direction, then it will be $k_2 = N - k_1$ steps away in the counter-clockwise direction. Each of the other nodes can thus always be reached by two complementary steps with sizes satisfying $k_1 + k_2 = N$. To avoid this ambiguity, we will use the smaller length for our steps. This choice also allows us to define the length of each edge as the magnitude of the corresponding step. We will also define a clockwise step as positive and a counter-clockwise step as negative. Finally, when N is even, the clockwise and counter-clockwise steps directly across the circle will both have length $N/2$ with opposite signs. This ambiguity corresponds exactly to the $\pm 180°$ ambiguity when measuring angles. To avoid this issue, we will simply define a step directly across the circle as positive $N/2$.

Putting all of these definitions together, we arrive at a unique set of $N-1$ possible steps s from a given node that satisfy $-N/2 < s \leqslant N/2$, with $s = N/2$ only allowed for even N and $s \neq 0$. For a path $\boldsymbol{p}$ with path differences $\boldsymbol{d}$ we can compute the vector of

[1]Notice $d_n \neq 0$ since all the $\{p_n\}_{n=0}^{N-1}$ are distinct.

DOI 10.2478/rmm-2025-0002

corresponding steps $\boldsymbol{s} \in \mathbb{Z}^N$ with entries $\{s_n\}_{n=0}^{N-1}$ as

$$s_n = \begin{cases} d_n, & |d_n| < \frac{N}{2} \\ \frac{N}{2}, & |d_n| = \frac{N}{2} \\ d_n - N, & d_n > \frac{N}{2} \\ d_n + N, & d_n < -\frac{N}{2} \end{cases} \tag{2}$$

Figure 4 illustrates these steps from the top node of the circle for $N = 7$ and $N = 8$. As an example, for the $N = 8$ asymmetric star `NITROGEN` $\rightarrow$ `RINGTONE` we obtain

$$\begin{array}{rcl} \boldsymbol{p} & = & [\ 3\ \ 1\ \ 7\ \ 5\ \ 2\ \ 4\ \ 0\ \ 6\ ] \\ \boldsymbol{d} & = & [\ \text{-}2\ \ 6\ \ \text{-}2\ \ \text{-}3\ \ 2\ \ \text{-}4\ \ 6\ \ \text{-}3\ ] \\ \boldsymbol{s} & = & [\ \text{-}2\ \ \text{-}2\ \ \text{-}2\ \ \text{-}3\ \ 2\ \ 4\ \ \text{-}2\ \ \text{-}3\ ]. \end{array}$$

This star is shown on the left of Figure 5 with the steps labeled.

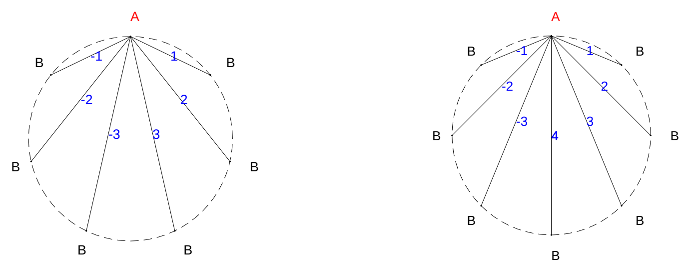

Figure 4: All possible steps from the red "A" node for $N = 7$ (left) and $N = 8$ (right).

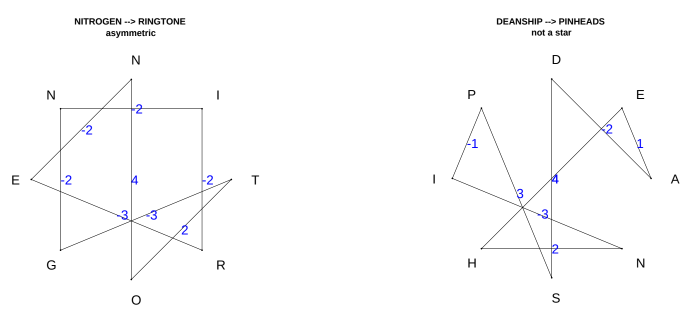

Figure 5: Two example anagrams with the steps $\{s_n\}_{n=0}^{N-1}$ labeled in blue.

Lemma 1 (Steps)**.** *The steps* $\{s_n\}_{n=0}^{N-1}$ *for a path* $\boldsymbol{p} \in \mathbb{Z}^N$ *satisfy*

$$p_{n+1} = (p_n + s_n) \bmod N \tag{3}$$

Proof. Examining (2), we see that $s_n \equiv d_n \pmod{N}$. Notice also that $p_n = p_n \bmod N$. Using (1), we combine these facts to obtain $p_{n+1} = p_{n+1} \bmod N = (p_n + d_n) \bmod N = (p_n + s_n) \bmod N$. $\square$

This relationship will be useful for proving properties of edge lengths in the sequel.

DOI 10.2478/rmm-2025-0002

2.4 Detecting a Star Path

Our remaining task for this section is to determine if a given path corresponds to a star anagram.

Theorem 1 (Star Detection). *A path $\boldsymbol{p} \in \mathbb{Z}^N$ is a star anagram path if and only if the path's steps satisfy $|s_n| \neq 1$ for all $n \in \{0, 1, \ldots N-1\}$.*

Proof. Recall that an anagram is a star if no letter in the new word retains its original neighbors. By (3), this condition occurs exactly when no step is to the nearest clockwise neighbor ($s_n = 1$) or the nearest counter-clockwise neighbor ($s_n = -1$). Recalling that $d_{N-1} = p_0 - p_{N-1}$, we see that testing the N steps captures all pairs of possible former neighbors. □

Notice that this check is easily performed for each of the P possible paths for each anagram, requiring only a few operations on N scalar values.

3 Star Anagram Classification

Now that we have a reliable method for detecting star anagrams, we turn our attention to classification. We start with the simpler test for perfection.

3.1 Identifying Perfection

Recall that a perfect star anagram has all edges of the same length. Since the edge lengths are given by the magnitudes of the steps $\boldsymbol{s}$, we arrive immediately at our test for perfection.

Theorem 2 (Perfection Test). *A path $\boldsymbol{p} \in \mathbb{Z}^N$ is a perfect star path if and only if the path's steps satisfy $s_n = S$ for all $n \in \{0, 1, \ldots N-1\}$ for a constant S satisfying $|S| = L > 1$.*

Proof. By definition, a perfect star path must satisfy $|s_n| = L$ for a constant $L > 1$. We see that the steps must have the same sign, since two consecutive steps with equal magnitude and opposed signs would cause the path to repeat the previous node. □

To see this test in action, consider our earlier example `EARTH` → `HATER`. We obtain

$$\begin{aligned} \boldsymbol{p} &= [\ \ 4\ \ 1\ \ \ 3\ \ 0\ \ 2\ ] \\ \boldsymbol{d} &= [\ \text{-}3\ \ 2\ \ \text{-}3\ \ 2\ \ 2\ ] \\ \boldsymbol{s} &= [\ \ 2\ \ 2\ \ \ 2\ \ 2\ \ 2\ ] \end{aligned}$$

This star, like all length 5 stars, is a perfect pentagram with $L = 2$. Indeed, this characteristic shape was the origin of the name star anagram.

3.2 Symmetry and Edges

Finally, our last task is to check if an anagram that is known to be a non-perfect star is symmetric. The key will again be to leverage our steps $\boldsymbol{s}$. However, we will find it convenient to work with a representation of the steps that makes the 2 edges associated with each node easier to analyze.

We define the matrix of edges $\boldsymbol{E} \in \mathbb{Z}^{2\times N}$ such that e_{1n} and e_{2n} are the two steps taken from the letter labeled n to the two letters connected to that node along the path. One of these steps is part of the path and corresponds exactly to s_k where k is the index for which $p_k = n$. The other is a step backwards along the path and will be $-s_{k-1}$ for the same k, where we define $s_{-1} = s_{N-1}$ for convenience. We also sort the entries such that $e_{1n} \leqslant e_{2n}$. Finally, we set steps across the circle with magnitude $N/2$ to 0 in the edge matrix to avoid ambiguities when dealing with negated values in the edge matrix.

An example will likely be helpful for the reader. Consider the non-star anagram `DEANSHIP` $\to$ `PINHEADS` shown on the right in Figure 5. We obtain the steps for this anagram as

$$\begin{aligned} \boldsymbol{p} &= [\ \ 7\ \ 6\ 3\ \ 5\ 1\ \ 2\ 0\ 4\ ] \\ \boldsymbol{d} &= [\ \text{-1}\ \text{-3}\ 2\ \text{-4}\ 1\ \text{-2}\ 4\ 3\ ] \\ \boldsymbol{s} &= [\ \text{-1}\ \text{-3}\ 2\ \ 4\ 1\ \text{-2}\ 4\ 3\ ] \end{aligned}$$

where $s_0 = -1$ and $s_4 = 1$ reveal the anagram to be a non-star. We obtain the edges $\boldsymbol{E}$ for this anagram as

$$\boldsymbol{E} = \begin{bmatrix} 0 & 0 & -2 & 2 & 0 & -2 & -3 & -3 \\ 2 & 1 & -1 & 3 & 3 & 0 & 1 & -1 \end{bmatrix}$$

Lemma 2. *Two paths produce identical shapes if and only if they have identical edge matrices.*

Proof. By construction, the edge matrix encodes the connections between nodes in our unicursal polygons. Sorting the entries within each column ensures a unique representation regardless of the direction or starting point of the paths. □

We say that two shapes are *equivalent* if they are identical up to a rotation in the plane. Noticing that rotating a shape does not change the lengths or signs of the edges, we see that rotating a shape in the plane is precisely equivalent to performing a circular shift of the columns in the edge matrix, proving the following.

Lemma 3. *Two shapes are equivalent if and only if their edge matrices are equal up to a circular shift of the columns.*

3.3 Identifying Rotational Symmetry

With the edge matrix $\boldsymbol{E}$ defined, we are ready to classify star anagram symmetry. We begin with the simpler case of rotational symmetry.

Consider rotating the shape created by drawing the star anagram's path completely through 360°. If throughout this single rotation, the anagram appears identical to its original position O_{rot} times, then we say that the anagram has rotational symmetry of order O_{rot}. Clearly, we have that $O_{rot} \leqslant N$.

Recalling Lemma 3, we can test for rotational symmetry by looking for periodicity in the columns of $\boldsymbol{E}$. This condition can be easily checked by circularly shifting the columns of $\boldsymbol{E}$ and testing for equality with the original matrix. We immediately see that if the smallest circular shift that produces the original matrix $\boldsymbol{E}$ is a shift of K, then the star anagram's shape will repeat N/K times through a complete rotation. Therefore, the anagram has rotational symmetry of order $O_{rot} = N/K$, with $O_{rot} = N/N = 1$ if it lacks rotational symmetry. We summarize these results in the following theorem.

Theorem 3 (Rotational Symmetry Test)**.** *A star anagram has rotational symmetry $O_{rot} = N/K$, where K is the smallest circular shift that recovers the original edge matrix.*

Notice that all the columns of $\boldsymbol{E}$ are identical for perfect star, yielding $O_{rot} = N$. Indeed, we note that order N rotational symmetry is an alternative test for perfection of a star anagram, excluding the non-star case with all $|s_n| = 1$ that produces a regular N-sided polygon.

Figure 6 provides examples of star anagrams with rotational symmetry of order one through three identified with this test. Other than perfect stars, we are unaware of a star anagram in English with $O_{rot} > 3$. Note that `INDENTING` $\rightarrow$ `INTENDING` also has reflective symmetry, which we will discuss next.

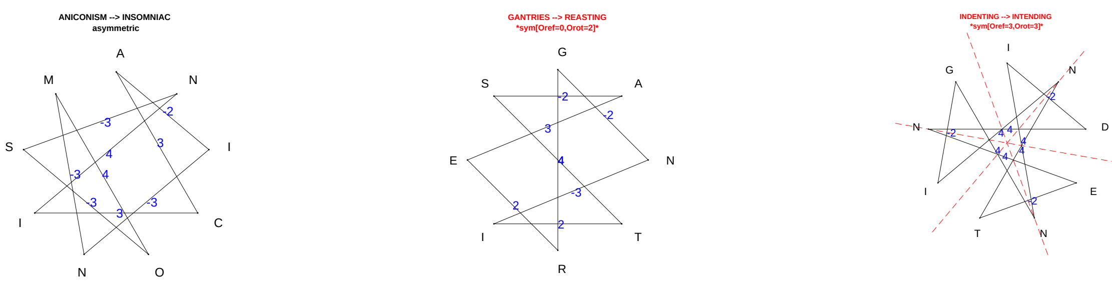

Figure 6: Examples of star anagrams with $O_{rot} = 1$ (left), $O_{rot} = 2$ (middle), and $O_{rot} = 3$ (right) with the $\{s_n\}_{n=0}^{N-1}$ labeled in blue.

3.4 Identifying Reflective Symmetry

A star anagram is said to be reflectively symmetric if the anagram's shape can be symmetrically folded across one or more dividing lines passing through the shape's center. We define the number of such lines to be the order of reflective symmetry O_{ref}. An asymmetric star anagram thus has $O_{ref} = 0$.

Such a line of symmetry can either pass through two nodes on the circle, pass between two nodes as it enters/leaves the circle, or pass through a node on one side and between two nodes on the other. Considering the geometry of these figures, the first and last cases can only happen when N is even, while the middle case always occurs when N is odd. Figure 7 provides examples of these 3 conditions to make the ideas concrete. Notice that all perfect stars satisfy $O_{ref} = N$, with a line of symmetry passing through every node (or pair of nodes for even N) and another $N/2$ lines of symmetry passing between nodes for even N. Similar to the rotational case, order N reflective symmetry is an alternative test for perfection, again excluding the regular polygons produced with all $|s_n| = 1$.

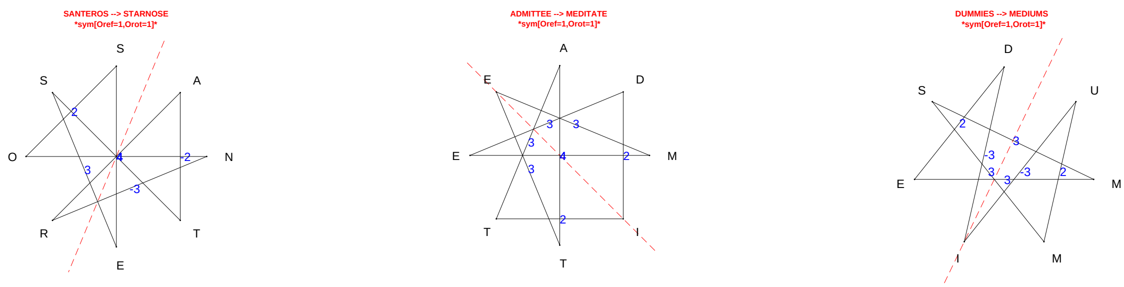

Figure 7: Examples of the 3 possible reflective symmetry conditions with the steps $\{s_n\}_{n=0}^{N-1}$ labeled in blue. (Left) N even, between nodes. (Middle) N even, through nodes. (Right) N odd, passing through one node.

Given the path $\boldsymbol{p}$ for a non-perfect star with even N, we need to look for lines of symmetry that pass either through or between nodes, as in the first two panels of Figure 7. We start with the more intuitive case of a line of symmetry that passes between nodes. Examining the figure, the condition we want to test is that paired nodes walking around the circle in opposite directions have edges with equal magnitudes and opposite signs, i.e., that are mirror images of each other.

In terms of the matrix $\boldsymbol{E}$, this situation will occur when the columns can be split into two groups with the second group in reverse order and negated. Note that the columns must be re-sorted after negation and reversing their order to test for equality. We refer to columns that can be split into two such groups as a *negated palindrome*. We can test our anagram for each of the $N/2$ possible lines of symmetry by circularly shifting $\boldsymbol{E}$ so that each of the first $N/2$ nodes are in the zeroth column and checking for this negated symmetry in the columns.

Testing for the second case in Figure 7 for even N is similar, except we have to handle the two nodes that the line of symmetry passes through differently. Rather than appearing as a pair of mirrored nodes, these nodes need to satisfy $e_{1n} = -e_{2n}$ so that the symmetry of the two edges "departing" from the node is maintained. The process for testing a star anagram with N odd is also similar, except each of the possible N lines of symmetry to be checked passes through exactly one node. We summarize these arguments with the following theorem.

Theorem 4 (Reflective Symmetry Test)**.** *A star anagram path has reflective symmetry order* $0 \leqslant O_{ref} \leqslant N$ *corresponding to the number of lines of reflective symmetry. Each* *possible line of symmetry is a line of reflective symmetry if and only if the edge matrix* $\boldsymbol{E}$ *satisfies* $e_{1n} = -e_{2n}$ *for each node* n *intersected by the line of symmetry, with the remaining columns forming a negated palindrome.*

- *For* N *even, there are* $N/2$ *possible lines of symmetry passing through opposing nodes and* $N/2$ *possible lines of symmetry passing between pairs of nodes.*
- *For* N *odd, each possible line of symmetry passes through exactly one node.*

3.5 Symmetry and Multiple Paths

We consider a star to be symmetric if it has rotational ($O_{rot} > 1$) or reflective ($O_{ref} > 0$) symmetry. Recall that we test each of the P possible paths $\{\boldsymbol{p}_i\}_{i=0}^{P-1}$ for an anagram. We still need to address how we choose the path to represent the star, denoted $\boldsymbol{p}^* \in \{\boldsymbol{p}_i\}_{i=0}^{P-1}$, when multiple paths for a given non-perfect star anagram are symmetric.

In most cases, one of the paths will have the sum $O_{rot} + O_{ref}$ greater than the sum for all other paths. In these cases, we simply take this strictly more symmetric, or *dominant*, path for the anagram as our selection $\boldsymbol{p}^*$ and report the order of its symmetries. An interesting example is the symmetric star `MOORWORT` $\rightarrow$ `ROOTWORM`. The 3 symmetric paths of its 12 possible paths are shown in Figure 8. The first path is only rotationally symmetric, the second is only reflectively symmetric, and the third path, which we select as $\boldsymbol{p}^*$, is both. There are also a handful of cases that have no clearly dominant path. As an example, consider the symmetric star `BORROWER` $\rightarrow$ `REBORROW`. Both of its symmetric paths shown in Figure 8 satisfy $O_{rot} + O_{ref} = 2$, and we simply select one arbitrarily to represent the anagram.

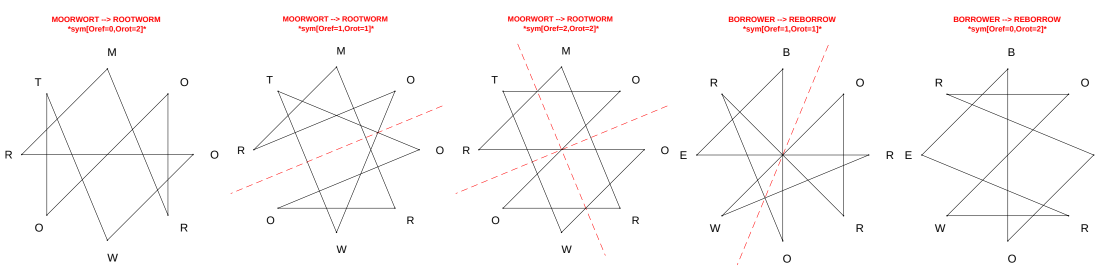

Figure 8: (Left 3 panes) The 3 symmetric star paths for the anagram `MOORWORT` $\rightarrow$ `ROOTWORM`. (Right 2 panes) The 2 symmetric star paths for the anagram `BORROWER` $\rightarrow$ `REBORROW`.

3.6 Summary of Star Anagram Detection and Classification

We briefly summarize our overall approach to star anagram detection and classification. First, for ease of discussion, we consider "non-star" as a fourth class, rather than describing detection as a separate process. Thus, we consider placing a given anagram into one of four classes: non-star, asymmetric star, symmetric star, and perfect star. We also consider the classes to be ordered as given so that we can apply comparison operators between classes when describing our procedure.

With these ordered classes in mind, we first generate the set of all possible paths for the anagram. Next, for each path, we determine the class and symmetry orders (O_{rot}, O_{ref}) as described in Procedure 1. Looking across these results, we select the path that gives the “best” or greatest class given our ordering above. Ties among paths of the same class are broken by maximizing the sum $O_{rot} + O_{ref}$. For paths with the same class and sum of symmetry orders, we simply select the last one in our enumerated list. This overall process is summarized in Procedure 2.

Procedure 1 Classify a path $\boldsymbol{p}$

Input: $\boldsymbol{p}$
Output: class $\in \{\textbf{non-star}, \textbf{asymmetric}, \textbf{symmetric}, \textbf{perfect}\}$, O_{rot}, O_{ref}

```
  Initialize:
  N ← length of p
  O_rot ← −1; O_ref ← −1 {Not determined}

  Classify:
  Compute s {Compute steps}
  if |s_n| = 1 for any n ∈ {0, 1, ... N − 1} then
    class ← non-star
    return {Not a star. Stop.}
  end if
  if |s_n| = L,   ∀n ∈ {0, 1, ... N − 1}, then
    class ← perfect
    O_rot ← N; O_ref ← N
    return {Perfect star found. Stop.}
  end if
  Compute E for p {Compute edges}
  Compute O_rot, O_ref for E {Test for symmetry}
  if O_rot + O_ref > 1 then
    class ← symmetric
  else
    class ← asymmetric
  end if
```

4 Remarks on Star Anagrams

Now that we can detect and classify star anagrams, we make some remarks about interesting properties of certain star anagram classes.

DOI 10.2478/rmm-2025-0002

Procedure 2 Classify the anagram $w_1 \to w_2$

Input: w_1, w_2

Output: $\text{class}^* \in \{\textbf{non-star}, \textbf{asymmetric}, \textbf{symmetric}, \textbf{perfect}\}$, $\boldsymbol{p}^*$, O^*_{rot}, O^*_{ref}

Initialize:

$\text{class}^* \leftarrow \textbf{non-star}$ {Define: **non-star** < **asymmetric** < **symmetric** < **perfect**}

$\boldsymbol{p}^* \leftarrow \boldsymbol{p}_1$

$O^*_{rot} \leftarrow -1$; $O^*_{ref} \leftarrow -1$ {Not determined}

Compute all paths $\{\boldsymbol{p}_i\}_{i=0}^{P-1}$ for $w_1 \to w_2$

Test each path:

for $i = 0$ **to** $P - 1$ **do**

Compute $(\text{class}^i, O^i_{rot}, O^i_{ref})$ for $\boldsymbol{p}_i$ {Using Procedure 1}

if $\text{class}^i > \text{class}^*$ **or** $(\text{class}^i = \text{class}^*$ **and** $O^i_{rot} + O^i_{ref} \geqslant O^*_{rot} + O^*_{ref})$ **then**

$\text{class}^* \leftarrow \text{class}^i$

$\boldsymbol{p}^* \leftarrow \boldsymbol{p}^i$

$O^*_{rot} \leftarrow O^i_{rot}$; $O^*_{ref} \leftarrow O^i_{ref}$

end if

end for

4.1 Perfect Star Edge Lengths

Recall that a perfect star is one for which $|s_n| = L$ for all n. One naturally asks: what edge lengths L are possible for perfect stars with a given word length $N \geqslant 5$?

First, we note that $L \geqslant 2$, since $L = 1$ does not produce a star path. Next, recalling the definition of the steps $\boldsymbol{s}$, we see that[2] $L < N/2$. As noted previously, one can easily show that the steps s_n for a perfect star all have the same sign. Since reversing the path for a perfect star will produce the same shape, we can without loss of generality consider only positive valued steps with $s_n = L$ for all n. Furthermore, since a perfect star path can be rotated to start with any node, we can without loss of generality consider paths that begin with $p_0 = 0$.

Our task is thus reduced to finding the constant step sizes $2 \leqslant L < N/2$ such that a path with these steps and $p_0 = 0$ forms a perfect star. Incrementing in constant steps of size $s_n = L$ and recalling equation 3, we obtain for $1 \leqslant n \leqslant N - 1$

$$p_n = (p_{n-1} + L) \bmod N = \left(p_0 + \sum_{i=0}^{n-1} L\right) \bmod N = q_n \bmod N$$

where we define $q_n = nL$ for $n = 0, 1, \ldots N-1$. Notice that we also have $p_0 = q_0 \bmod N$, since $q_0 = 0$. We also state a simplified version of the Chinese Remainder Theorem [Ros11, Theorem 4.13] that will be useful in our proof.

[2]We will soon see that $L = N/2$ is not a valid edge length for a perfect star; hence the strict inequality here.

DOI 10.2478/rmm-2025-0002

Theorem 5 (Simplified Chinese Remainder Theorem). *For coprime integers $A, B > 0$, each integer $C \geqslant 0$ with $C < AB$ can be uniquely identified from the remainders $C \bmod A$ and $C \bmod B$.*

With these ideas in mind, we can arrive at the following.

Theorem 6 (Valid Edge Lengths for Perfect Stars). *A perfect star path with length N and constant edge length L exists if and only if L and N are coprime with $2 \leqslant L < N/2$ and $N \geqslant 5$.*

Proof. First, we show that an L that is not coprime with N cannot be a valid edge length for a perfect star. Assume L and N are *not* coprime. Then there exists integers $a, b \geqslant 2$ and $c \geqslant 1$ such that $N = ab$ and $L = ac$. We see that $p_b = bL \bmod N = bac \bmod N = Nc \bmod N = 0$. Since $b < N - 1$, this implies that the path revisits node 0 before visiting all the other nodes. Thus, this is not a valid path for a star anagram, and L is not a valid edge length for a perfect star. Note that this also excludes $L = N/2$ as a valid edge length, since $N/2$ and N are not coprime.

Now assume that L and N are coprime. Notice that $q_n < NL$ for all n. Since N and L are coprime, Theorem 5 requires that any integer $q_n < NL$ be uniquely identifiable from the values $q_n \bmod L$ and $q_n \bmod N$. Since, $q_n \bmod L = nL \bmod L = 0$ for all n, this requires that $q_n \bmod N$ is distinct for each n. But these N values between 0 and $N - 1$ are precisely the elements $\{p_n\}_{n=0}^{N-1}$, which shows that this path visits each node exactly once. Thus, this path forms a valid perfect star path with $s_n = L$. $\square$

We note that this result is already well known in the study of star polygons, e.g., [Cox69, Section 2.8].

Notice the interesting consequence that there are no perfect stars for words of length $N = 6$. We note that this result is perhaps ironic given that 6 is a perfect integer[3]. While perfect stars of any other length are possible, we will see below that perfect star anagrams for large N were not found in our search over the English language.

4.2 Commutativity

We defined an anagram to be an ordered pair. A natural question arises: if $w_0 \to w_1$ is a star, does it follow that $w_1 \to w_0$ is also a star? Put another way, is starriness commutative? We shall see that this is indeed the case.

Let $\boldsymbol{p}$ with entries $\{p_n\}_{n=0}^{N-1}$ be a path for $w_0 \to w_1$. We wish to construct the reversed path[4] $\bar{\boldsymbol{p}}$ with entries $\{\bar{p}_n\}_{n=0}^{N-1}$ for $w_1 \to w_0$. If we label each letter with its position in both words, we notice that the two sets of values exchange roles as the indices and node locations for the path and its reversed path. An example may be helpful; consider our

[3]The integer 6 is prefect in the sense that $6 = 1 + 2 + 3$. In other words, 6 is equal to the sum of its factors.

[4]Note that each path for an anagram with repeated letters has a corresponding reversed path.

familiar $N = 5$ star with the paths labeled in both directions

$$\begin{array}{cccccccccccc} & & & & & & p_0 & p_1 & p_2 & p_3 & p_4 \\ 0 & 1 & 2 & 3 & 4 & \rightarrow & 4 & 1 & 3 & 0 & 2 \\ \mathrm{E} & \mathrm{A} & \mathrm{R} & \mathrm{T} & \mathrm{H} & & \mathrm{H} & \mathrm{A} & \mathrm{T} & \mathrm{E} & \mathrm{R} \\ 3 & 1 & 4 & 2 & 0 & \leftarrow & 0 & 1 & 2 & 3 & 4 \\ \bar{p}_0 & \bar{p}_1 & \bar{p}_2 & \bar{p}_3 & \bar{p}_4. & & & & & & \end{array}$$

Notice that the pair of values above and below each letter remains the same in both words. In general, we see that the two paths satisfy

$$p_{\bar{p}_n} = n \tag{4}$$
$$\bar{p}_{p_n} = n \tag{5}$$

for $n = 0, 1, 2, \ldots N-1$. This relationship will be useful shortly when we examine perfect stars. We first state a result on general star anagrams.

Theorem 7 (Starriness Commutes)**.** *If $w_0 \rightarrow w_1$ is a star anagram with star path $\boldsymbol{p}$, then $w_1 \rightarrow w_0$ is a star anagram with the reversed path $\bar{\boldsymbol{p}}$.*

Proof. Suppose that $\bar{\boldsymbol{p}}$ is not a star path. This requires a pair of letters that are adjacent in w_0 to also be adjacent in w_1. However, this contradicts the starriness of $\boldsymbol{p}$, and we conclude that $\bar{\boldsymbol{p}}$ is a star path for $w_1 \rightarrow w_0$. □

Next we turn our attention to perfection, which also commutes. However, the edge length of the reversed star is often not the same. An example is provided in Figure 9 with $L = 3$ for the perfect star and $L = 2$ for the reversed perfect star. The following theorem provides the exact relationship between the edge lengths in terms of the possibly negative constant step size S, recalling that $L = |S|$ for a perfect star.

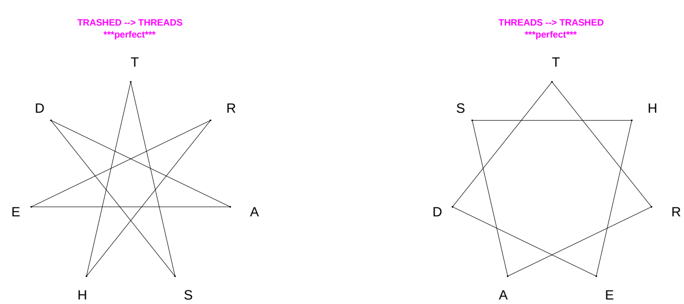

Figure 9: An example perfect star with $S = -3$ (left) and the corresponding reversed perfect star (right) with $\bar{S} = 2$. As expected from Theorem 8, $\bar{S}S \bmod N = 2(-3) \bmod 7 = 1$.

Theorem 8 (Perfection Commutes)**.** *If $w_0 \rightarrow w_1$ is a length N perfect star anagram with path $\boldsymbol{p}$ and constant step size S, then $w_1 \rightarrow w_0$ is a prefect star with the reversed path $\bar{\boldsymbol{p}}$ and step size $\bar{S}$ satisfying $\bar{S}S \bmod N = 1$.*

DOI 10.2478/rmm-2025-0002

Proof. From (3), $\boldsymbol{p}$ satisfies $p_n = (p_0 + nS) \bmod N$ for $n = 0, 1, 2, \ldots N-1$. Since this a star path, there must exist a unique $2 \leqslant K < N$ such that $p_K = (p_0 + KS) \bmod N = (1 + p_0) \bmod N$, i.e., the path must visit the node 1 greater than p_0 exactly once and not on the step p_1. Adding the integer $p_n - p_0$ to both sides, we obtain for each n and our constant K the relationship $(1 + p_n) \bmod N = (p_n + KS) \bmod N = p_{(n+K \bmod N)}$, where the last equality follows by noticing that $p_0 = (p_0 + NS) \bmod N$ and the properties of modular arithmetic. Note that $p_n = 0$ for some n, and hence $KS \bmod N = 1$.

Consider now the path differences for the reversed path $\{\bar{d}_n\}_{n=0}^{N-1}$. By definition we have $\bar{d}_n = \bar{p}_{n+1} - \bar{p}_n$. Recalling that we defined $p_N = p_0$ for notational convenience, we can more precisely write this as $\bar{d}_n = \bar{p}_{(n+1 \bmod N)} - \bar{p}_n$. For a given n, recalling equation 4, there is a unique n^* such that $p_{n^*} = n$. From our results above, we have that $(n+1) \bmod N = (p_{n^*} + 1) \bmod N = p_{(n^*+K \bmod N)}$. Putting these together and applying (5), this allows us to obtain $\bar{d}_n = \bar{p}_{p_{(n^*+K \bmod N)}} - \bar{p}_{p_{n^*}} = (n^* + K) \bmod N - n^*$. Since $0 \leqslant n^*, K < N$, we can easily see that $\bar{d}_n \in \{K, K - N\}$ for $n = 0, 1, 2, \ldots N-1$.

Recalling the definition of $\bar{s}_n$ from (2), we can see that $\bar{s}_n = K$ for all n if $K \leqslant N/2$ or $\bar{s}_n = K - N$ for all n if $K > N/2$. Thus, $\bar{\boldsymbol{p}}$ has a constant step step size $\bar{S} \in \{K, K-N\}$ and is a perfect star path. Recalling $KS \bmod N = 1$, and adding $-NS$ if $\bar{S} = K - N$, we have $\bar{S}S \bmod N = 1$. $\square$

We note that $\bar{S}$ is the *modular inverse* of S in the parlance of number theory, which exists uniquely modulo N precisely when N and S are coprime [Ros11, Section 4.2]. Our perfect step size S is of course coprime with N by Theorem 6.

Finally, we address the reversed paths of symmetric stars. It is not the case that reversing the path preserves rotational or reflective symmetry. A counter example is the symmetric star `TOENAILS` $\to$ `INSOLATE` shown in Figure 10. However, we have verified by brute force that the reversed star of every symmetric star has some form of symmetry, which suggests the conjecture that general symmetry commutes. We defer a rigorous analysis and possible proof of this conjecture to future work.

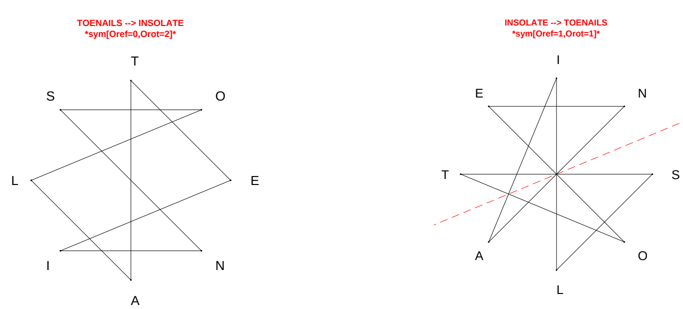

Figure 10: An example symmetric star with a reversed path that differs in both symmetry orders.

DOI 10.2478/rmm-2025-0002

4.3 Star Clusters

Recall that two shapes are equivalent if they are identical after a rotation in the plane. We will refer to a set of star anagrams with equivalent polygons as a *star cluster*. From Theorem 6, we know that the number of possible perfect star polygons for length N is simply the number of integers less than $N/2$ that are coprime with N. The total number of possible star polygons was calculated in [Gor13], specializing earlier work in [GW60] on the enumeration of polygons in general.

We provide the number of possible star polygons of each class for several lengths N in Table 1. The counts in the table were computed by brute force enumeration and match the theoretical results for the perfect and total columns. We are not aware of a closed form result for the number of possible symmetric star polygons and leave such a result for future work.

	Unique star polygons up to a rotation			
N	Asymmetric	Symmetric	Perfect	Total
5	0	0	1	1
6	0	1	0	1
7	0	3	2	5
8	12	14	1	27
9	126	47	2	175
10	1354	178	1	1533

Table 1: Number of possible unique star polygons, up to a rotation, of each class for several word lengths N. The perfect column is consistent with Theorem 6, and the total column matches the results in [Gor13].

4.4 Autostars

For a word with no repeated letters, the only path that produces the original word will be a regular polygon, making a simple trip around the circle. This does not meet our definition of an anagram, since the letters are not re-arranged. We might denote this trivial case as a *self anagram*. However, a word with repeated letters admits multiple paths around the circle that produce the original word. It turns out that some words with repeated letters are thus star anagrams of themselves. We refer to these words as *autostars*. In the following section, we will see that autostars are quite common in English and provide interesting star shapes not observed among normal star anagrams.

5 Star Search

We now turn our attention toward detecting and classifying all star anagrams in a set of $501,603$ English words[5]. Particularly for large N, this extensive word list includes many jargon and highly obscure words.

Our code first scans this word list and produces an exhaustive enumeration of all possible anagrams, yielding $108,716$ total anagrams. Each anagram is then processed using the methods in Section 3 to classify it as a non-star, asymmetric star, symmetric star, or perfect star. The resulting collection of star anagrams is organized first by length N, then by class, and finally by shape into star clusters. This process produced a total of $6,204$ star anagrams ranging in length from $N = 5$ to $N = 21$. These results included $1,614$ asymmetric stars, $3,335$ symmetric stars, and $1,255$ perfect stars. Thus, about 5.7% of English anagrams are stars.

Figure 11 summarizes the results. Each vertical bar shows the number of stars found for a given value of N, broken into up to 3 pieces to show the split between asymmetric, symmetric, and perfect stars for that word length. Only lengths 8, 9, and 10 have stars from all three classes, and hence only those bars are split into 3 pieces. The black number on each bar piece indicates the number of star clusters found for that class and word length. Comparison with Table 1 reveals that all possible shapes are only observed for $N < 9$.

We provide examples from each of the symmetric and perfect clusters for $N \in \{6, 7, 8\}$ in Figure 12. We also include either common names for these shapes or our own nicknames. The perfect stars are denoted with their key parameters written as $\{N/L\}$, a convention known as *Schläfli notation* [Cox69, Section 2.8]. Examples of the longest stars found for each class are shown in Figure 13. Star anagrams become exceedingly rare for large N. Indeed, we found only 25 stars with $N > 14$. A complete set of figures depicting all the star anagrams found in our search is available online [PB22].

5.1 Autostar Results

Words that are autostars are surprisingly common, with $9,948$ examples found in our search over the English language. For very long words with repeated letters, the number of possible paths to search becomes computationally prohibitive. For example, the word SUPERCALIFRAGILISTICEXPIALIDOCIOUS has $209,018,880$ possible autostar paths. To avoid this problem, we excluded words with more than $3,000,000$ possible paths from our search for autostars. Only 21 words were eliminated by this criteria from our list of $501,603$ English words.

Autostars can of course be symmetric. We found $3,542$ symmetric autostars. Two interesting examples are shown in Figure 14. The first example in the figure is particularly interesting, since it has rotational symmetry of order 4 without being perfect. We are

[5]We used the Spell Checker Oriented Word List (SCOWL) [Atk18] Version 2020.12.07 with United States spelling including British variations, “insane” (95) size, common spelling variants to level 3, and included “hacker” words in the list. We also note that accent marks were suppressed for anagram search purposes.

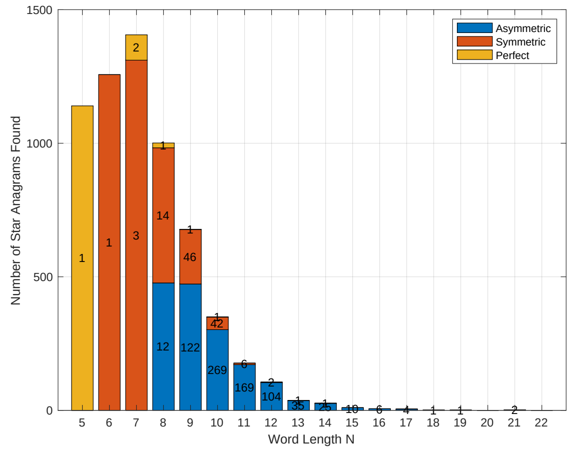

Figure 11: The number of each class of star anagram found for each value of word length N. The black number on each bar indicates how many unique shapes were found for that class and word length.

not aware of any regular stars with this property. The longest symmetric autostar we found has length $N = 20$ and appears in the Figure's second pane. We also found perfect autostars with length 12, exceeding the maximum length 10 we found among regular stars. The third pane of Figure 14 provides an example. Finally, the last two panes of Figure 14 show two perfect paths for the autostar `BERSERKER` $\rightarrow$ `BERSERKER` with different edge lengths L. We did not find such an example among regular stars.

6 Conclusions

We have developed an efficient algorithmic approach for detecting and classifying star anagrams and proved several interesting properties relating to their edge lengths and reversability. While we proved that starriness and perfection are preserved by reversing a star anagram, we leave for possible future work a proof that symmetry is preserved after reversal.

Results from a detailed numerical study of all star anagrams in a large database of English words were also presented, including geometrically interesting examples of autostars with properties not found among normal star anagrams. Results on the clustering of star anagrams into common shapes were also provided. Future work could include a rigorous analysis of the number of possible shapes of symmetric stars. Finally, a numerical study of star anagrams in other languages, including those with different alphabets, might yield interesting geometric results with different distributions across symmetry orders and word lengths than those found in English.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data is publicly available.

Acknowledgment

The authors conducted this work independently without sponsorship. The views in this article are not those of the U.S. government, the U.S. Air Force Research Laboratory, or the Freedom From Religion Foundation.

References

[Atk18] Kevin Atkinson. Spell checking oriented word list, 2018. URL: `wordlist.aspell.net` [cited 17 April 2022].

[Bar94] Daniel Barker. Star anagrams. *Gift of Fire*, (64):5–8, March 1994. URL: `https://wordsmith.org/anagram/star_anagrams.pdf`.

[Cox69] H. S. M. Coxeter. *Introduction to Geometry.* John Wiley & Sons, New York, London, Sydney, Toronto, 2nd edition, 1969.

[Gor13] S. Gordon. The online encyclopedia of integer sequences: Sequence a231091, 2013. URL: `oeis.org/A231091` [cited 9 September 2022].

[GW60] S. W. Golomb and L. R. Welch. On the enumeration of polygons. *American Mathematical Monthly*, 67(4):349–353, 1960.

[PB22] Jason Parker and Dan Barker. Star anagram detection and classification, 2022. URL: `https://arxiv.org/abs/2210.06397`, `doi:10.48550/ARXIV.2210.06397`.

[Ros11] K. H. Rose. *Elementary Number Theory and Its Applications.* Addison-Wesley, Reading, MA, 6th edition, 2011.

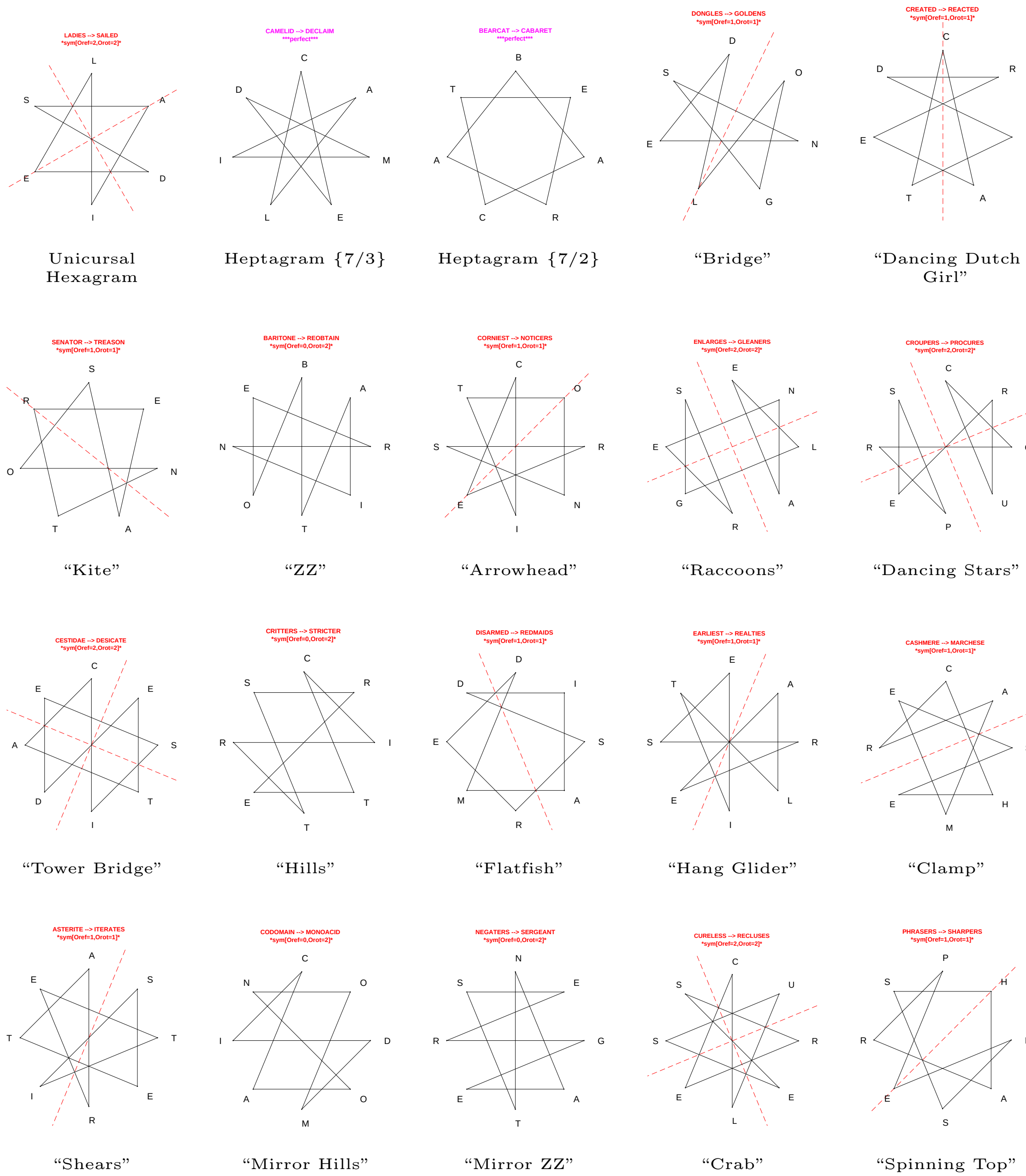

Figure 12: Examples of 20 of the 21 unique shapes for symmetric and perfect stars of length $N \in \{6, 7, 8\}$ along with our nicknames or common names for these shapes. See Figure 2 for an example Octagram $\{8/3\}$.

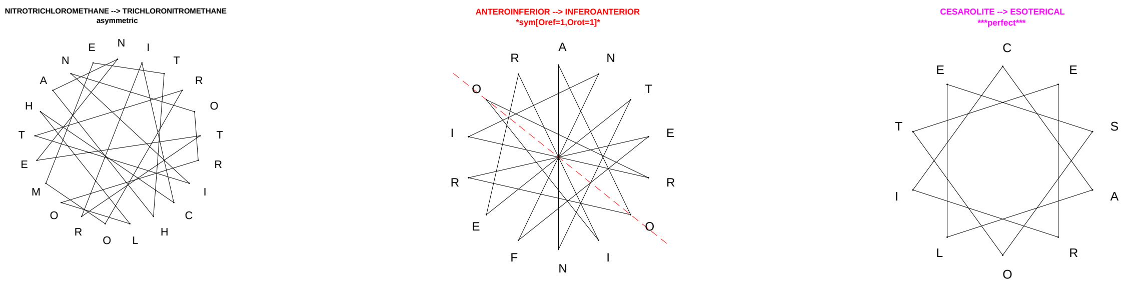

Figure 13: Examples of the longest stars found for each class. (Left) Asymmetric star with $N = 21$. (Middle) Symmetric star with $N = 14$. (Right) Perfect star with $N = 10$.

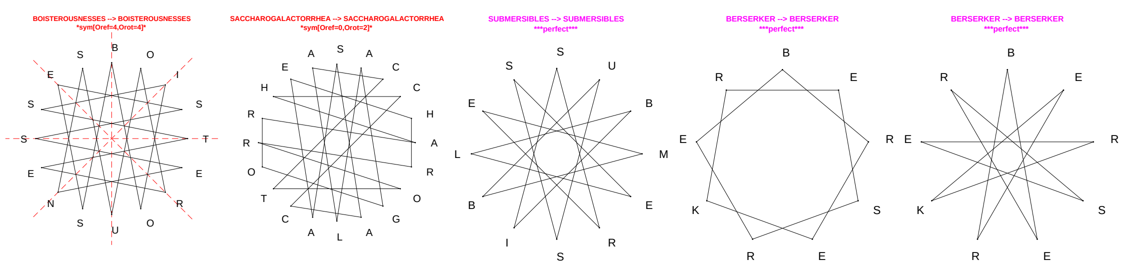

Figure 14: Example autostar results including a symmetric autostar with $O_{ref} = 4$, a symmetric autostar with $N = 20$, a perfect autostar with $N = 12$, and a perfect autostar with two possible edge lengths.