Counting the Number of Langford Skolem Pairings

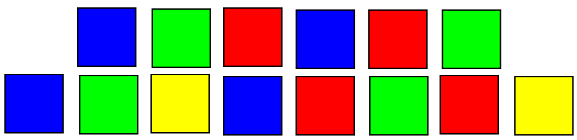
Ali Assarpour   Amotz Bar-Noy   Ou Liu

Abstract

We compute the exact number, $L(n)$, of solutions to the Langford pairings problem for any positive integer $n < 29$ and the exact number of solutions to the Nickerson variant of the problem, $N(n)$, for any positive integer $n < 26$. These numbers correspond to the sequences $A014552$, $A059106$ in Sloane's Online Encyclopedia of Integer Sequences. The exact value of these numbers were known for any positive integer $n < 27$ for the $A014552$ sequence and for any positive integer $n < 24$ for the $A059106$ sequence. First we report that the number of Langford pairings for $n = 27$ is $L(27) = 111,683,611,098,764,903,232$, and for $n = 28$ it is $L(28) = 1,607,383,260,609,382,393,152$. Next we report that the number of solutions for the Nickerson variant of Langford pairings for $n = 24$ is $N(24) = 102,388,058,845,620,672$ and for $n = 25$ it is $N(25) = 1,317,281,759,888,482,688$.

1 Historical Background

The Langford problem is named after C. D. Langford, a Scottish mathematician, who devised the problem in 1958 after having observed his son playing with colored blocks. He noticed that his son had arranged a set of six block, two red, two blue, and two green, in such away that the pair of red blocks were separated by a single block, the blue pair by two blocks, and the green pair by three blocks. He further noticed that he could add a pair of yellow blocks to the arrangement in away that would preserve the distances of the previous blocks while having the yellow pair separated by four blocks.



He captured this idea using numbers and asked the question, Given a sequence of $2n$ numbers $\{1, 1, 2, 2, \ldots, n, n\}$ find a permutation in which the two copies of each number k are k units apart. For instance, when $n = 4$ the permutation $2, 3, 4, 2, 1, 3, 1, 4$ would be a solution.
At about the same time while working on Steiner triple systems, Norwegian mathematician T. H. Skolem, in 1957, proposed a similar problem, he asked if it was possible to distribute the number $\{1, 2, \ldots, 2n\}$ in n pairs (a_r, b_r) such

that $b_r - a_r = r$ for $r = 1, 2, \ldots, n$?". For instance when $n = 4$, the pairings $(1, 4), (2, 6), (3, 5), (7, 8)$ is one solution.
In 1966 Nickerson, unaware of Skolem's problem, proposed a varient of Langford's problem where the pair of numbers k are separated by exactly $k-1$ other numbers. For instance when $n = 4$ the sequence $3, 4, 2, 3, 2, 4, 1, 1$ is a solution. Notice that if we consider the sequence as being placed in an array with $2n$ positions and index from 1 to $2n$ then this sequence is the same as the Skolem pairing.

1	2	3	4	5	6	7	8
3	4	2	3	2	4	1	1

The first natural question to asked is when do such pairings exist, that what are the necessary and sufficient conditions for the existance such sequences. In 1959 Daveis answers this question for the Langford formulation, and showed the sequence exists whenever $n \equiv 0, 3 \mod 4$, following the observation that that the pairs of number occupy the positions 1 to $2n$ in some order, the first occurrence of k is at position a_k and the second at $a_k + k + 1$. Furthermore he gave a method of construction such a sequence.

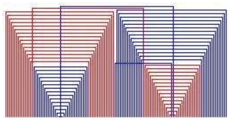
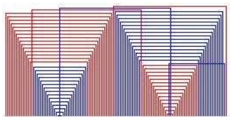
$n = 4k - 1$	4k-4, ..., 2k, 4k-2, 2k-3, ..., 1, 4k-1, 1, ..., 2k-3, 2k, ..., 4k-4, 2k-1, 4k-3, ..., 2k+1, 4k-2, 2k-2, ..., 2, 2k-1, 4k-1, 2, ..., 2k-2, 2k+1, ..., 4k-3	
$n = 4k$	4k-4, ..., 2k, 4k-2, 2k-3, ..., 1, 4k-1, 1, ..., 2k-3, 2k, ..., 4k-4, 4k, 4k-3, ..., 2k+1, 4k-2, 2k-2, ..., 2, 2k-1, 4k-1, 2, ..., 2k-2, 2k+1, ..., 4k-3, 2k-1, 4k	

Figure 1: Davies construction of a solution to the Langford pairing

In the case of the Skolem problem the existance question was answered by Skolem himself in 1957. He showed that the pairings exists whenever $n \equiv 0, 1 \mod 4$, following a similar observation that pairs of number occupy the positions 1 to $2n$ in some order, the first occurrence of k is at position a_k and the second at $a_k + k$. And a method of construction such a sequence was presented as well.

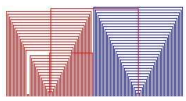
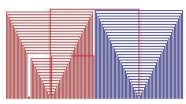
$n = 4k$	$(4k + r, 8k - r)$ for $r = 0, 1, \ldots, 2k - 1$, $(2k + 1, 6k)$, $(2k, 4k - 1)$, $(r, 4k - 1 - r)$ for $r = 1, 2, \ldots, k - 1$, $(k, k + 1)$, $(k + 2 + r, 3k - 1 - r)$ for $0, 1, \ldots, k - 3$.	
$n = 4k + 1$	$(4k + 2 + r, 8k + 2 - r)$ for $r = 0, 1, \ldots, 2k - 1$, $(2k + 1, 6k + 2)$, $(2k + 2, 4k + 1)$, $(r, 4k + 1 - r)$ for $r = 1, 2, \ldots, k$, $(k + 1, k + 2)$, $(k + 2 + r, 3k + 1 - r)$ for $r = 1, 2, \ldots, k - 2$.	

Figure 2: Skolem's construction of a solution to the Skolem pairing

n	$A014552$	$A059106$	$A176127$	$A004075$
1	0	1	0	1
2	0	0	0	0
3	1	0	2	0
4	1	3	2	6
5	0	5	0	10
6	0	0	0	0
7	26	0	52	0
8	150	252	300	504
9	0	1,328	0	2,656
10	0	0	0	0
11	17,792	0	35,584	0
12	108,144	227,968	216,288	455,936
13	0	1,520,280	0	3,040,560
14	0	0	0	0
15	39,809,640	0	79,619,280	0
16	326,721,800	700,078,384	653,443,600	1,400,156,768
17	0	6,124,491,248	0	12,248,982,496
18	0	0	0	0
19	256,814,891,280	0	513,629,782,560	0
20	2,636,337,861,200	5,717,789,399,488	5,272,675,722,400	11,435,578,798,976
21	0	61,782,464,083,584	0	123,564,928,167,168
22	0	0	0	0
23	3,799,455,942,515,488	0	7,598,911,885,030,976	0
24	46,845,158,056,515,936	102,388,058,845,620,672	93,690,316,113,031,872	204,776,117,691,241,344
25	0	1,317,281,759,888,482,688	0	2,634,563,519,776,965,376
26	0	0	0	0
27	111,683,611,098,764,903,232*	0	223,367,222,197,529,806,464*	0
28	1,607,383,260,609,382,393,152	???	3,214,766,521,218,764,786,304	???

Table 1: Langford number $A014552$, Nickerson's variant $A059106$, $A014552 = 2$ x $A014552$, and Skolem number $A059106 = 2$ x $A059106$ for $n > 1$, * $n = 27$ was previously reported to be $111,683,606,778,027,803,456$

In this note we are concered with the problem of counting the number solution. In what follows we present a breif overview of algorithms for counting these sequences, implementation and new results we found.

2 Counting Algorithms

There are in general three approaches to counting the number of pairings. First the generate and count methods, second the algebraic method, and third a method based on counting technique of inclusion exclusion. All the procedures below will be given interms of counting the number of Langford pairing, counting all solutions where reflected solutions are counted as distinct.

2.1 Miller's Generate and Count

One method of counting is to generate all possible permutations of the $2n$ numbers in some order and count as each valid solution is generated. However this is quit inefficient since there are $\frac{(2n)!}{2^n}$ permutations. A more efficient procedure is to use a backtracking algorithm for generating the solutions in a systematic way and counting them. One such procedure is given by Miller on his website [1], and reproduced below. The alorithm proceeds by placing the pairs in decreasing order, starting with leftmost available position where the pair can fit into. Once the pair is place it will try to palce the next smaller pair if it can not be placed then the previously placed pair needs to be moved to the next available valid position. The algorithm stops when all possible positions for the largest pair have been explored.

- Let n be the number of pairs to be placed

- Let $\mathcal{A}$ be an array with $2n$ positions holding the current (partial) arrangment
- Let k keep track the pair of numbers that are to be placed into $\mathcal{A}$
- Start by inserting the highest pair ($k = n$) into the first position. The pair will fit because the arrangement is empty.
- Next, decrement k and find the left most position p where both $\mathcal{A}[p]$ and $\mathcal{A}[p + k + 1]$ are empty
- If such a position p is found, then let $\mathcal{A}[p] = \mathcal{A}[p + k + 1] = k$.
  - If the 1's were just placed, a solution is reached, count it, and remove the 1's, 2's and 3's from $\mathcal{A}$, and continue scanning for the next available placment of $k = 3$.
- If the k^{th} pair can not be placed, forget about placing them, and increment the value of k. remove k's from $\mathcal{A}$, and find a new placement for the k^{th} pair starting from the next position to the right.
  - If k is incremented to beyond n, then the all the solutions have been found and algorithm terminates

2.2 Godfrey's Algebraic Method

An algebraic method for computing the number Langford pairings was proposed by Godfrey. There is no official paper on the algorithm but the description of this method appears in [9]. The problem is modeled by a polynomial where each term represents the label and its position, and the number of pairings is then the coefficient of the term $x_1 x_2 \dots x_n$. Consider $L(3)$, and $X = (x_1, x_2, x_3, x_4, x_5, x_6)$, then

$$F(3, X) = (x_1x_3 + x_2x_4 + x_3x_5 + x_4x_6)(x_1x_4 + x_2x_5 + x_3x_6)(x_1x_5 + x_2x_6)$$

each of the factors represents the possible ways in which a label and position can appear in the solution. For instance 3 can appear either in the first and fifth positions or in the second and sixth. When the polynomial is expanded then for example a term like $(x_2x_4)(x_3x_6)(x_1x_5)$ correspond to a sequence $3, 1, 2, 1, 3, 2$, and the coefficient to the term $x_1x_2x_3x_4x_5x_6$ will be the number of possible pairings (twice that since symmetric solutions are considered the same).

$$\begin{aligned} F(3, X) = {} & x_1^3x_3x_4x_5 + x_2^3x_4x_5x_6 + x_1x_2x_3x_5^3 + x_2x_3x_4x_6^3 + x_1^2x_3^2x_5x_6 + x_1^2x_2x_4^2x_5 + x_1^2x_4^2x_5x_6 + x_1^2x_2x_3x_5^2 + \\ & x_1^2x_3x_4x_5^2 + x_1x_2^2x_4^2x_6 + x_1x_2^2x_4x_5^2 + x_2^2x_3x_5^2x_6 + x_2^2x_3x_4x_6^2 + x_2^2x_4x_5x_6^2 + x_1x_3^2x_5^2x_6 + x_1x_2x_3^2x_6^2 + \\ & x_2x_3^2x_5x_6^2 + x_1x_2x_4^2x_6^2 + x_1^2x_2x_3x_4x_6 + x_1x_2^2x_3x_5x_6 + x_1x_2x_4x_5^2x_6 + x_1x_3x_4x_5x_6^2 + 2\, x_1x_2x_3x_4x_5x_6 \end{aligned}$$

However computing the coefficient by expanding the polynomial is just as time consuming as the generate and count methods since there are $\frac{(2n-2)!}{(n-2)!}$. To over come this the variables $x_1, \dots, x_{2n}$ are allowed to take on the values 1 and -1, and the resulting values of $(x_1 \times \dots \times x_{2n}) \times F(X, n)$ are summed, it is easy to see that the result is $2^{2n}L(n)$ (twice that incase symmetric solutions are counted as the same)

$$\sum_{(x_1,\dots,x_{2n})\in\{1,-1\}}\left(\prod_{i=1}^{2n}x_i\right)\prod_{i=1}^{n}\sum_{k=1}^{2n-i-1}x_kx_{k+i+1}=2^{2n}L(n)$$

Each of the terms in F other than $x_1x_2\dots x_{2n}$ is missing at least one of the variables (x_k), then these terms when multiplied by x_k result in zero after summing over $x_k=+1$ and -1. The calculation costs 2^{2n} evaluations of $F(n,X)$, each evaluation takes $O(n^2)$ multiplications and additions.

In practice the cost of multiplications can be reduced to $O(n)$ by going through the $\{-1,1\}^{2n}$ strings in Gray code order, there by replacing the multiplication by an addition and subtraction.

2.3 Larsen's Inclusion Exclusion Method

In 2009 Larsen [8] proposed a new algorithm based on the principle of inclusion exclusion to count the number of pairings. Inclusion exclusion is a counting technique for obtaining the number of elements in the union of sets. In its general form the principle of inclusion exclusion for finite sets $A_1,\dots,A_n$ states that the number of elements in the union of the sets is the sum of cardinalities of set intersections, subtracting when the number of sets is even and adding when its odd.

$$\left|\bigcup_{i=1}^{n}A_i\right|=\sum_{I\subseteq\{1,\dots,n\},I\neq\emptyset}(-1)^{|I|-1}\left|\bigcap_{i\in I}A_i\right|$$

It is common to generalize this form to count the number of elements that are not in any of the subsets of A_i. This complementary form is obtained by letting S be a finite universal set containing all of the A_i and letting $\overline{A_i}$ denote the complement of A_i in S, then

$$\left|\overline{\bigcup_{i=1}^{n}A_i}\right|=\left|\bigcap_{i=1}^{n}\overline{A_i}\right|=\left|S-\bigcup_{i=1}^{n}A_i\right|$$

or

$$\left|\bigcap_{i=1}^{n}\overline{A_i}\right|=\sum_{I\subseteq\{1,\dots,n\}}(-1)^{|I|}\left|\bigcap_{i\in I}A_i\right|$$

where $A_\emptyset=S$ since intersecting with no sets will result in all the elements.

This is connected to the number of pairings by viewing each set A_i as the set of (invalid) sequences that avoid the position i. Then $\overline{A_1}\cap\dots\cap\overline{A_n}$ is the set of sequences avoid no positions, and a valid solution to the problem. Then $L(n)$ can ne expressed in the follwoing form

$$L(n)=\sum_{X\subseteq\{1,\dots,2n\}}(-1)^{|X|}a(X)$$

where X denotes the positions to be avoided and $a(X)$ is the function that counts the set of solutions that avoid the positions in X. Let X be a binary sequence $(x_1,\dots,x_{2n})\in\{0,1\}$ where every 0 in the sequence is a disallowed poisition. Then

$$a(X)=\prod_{k=1}^{n}\sum_{j=1}^{2n-k-1}x_jx_{j+k+1}$$

To compute $L(n)$ need to sum over all binary strings of length $2n$.

$$L(n) = \sum_{(x_1,\ldots,x_{2n})\in\{0,1\}} \left((-1)^{\sum_{i=1}^{2n} x_i}\right) \prod_{k=1}^{n} \sum_{j=1}^{2n-k-1} x_j x_{j+k+1}$$

There will be a total 2^{2n} computations of $a(X)$, where each computation costs $O(n^2)$ time. The total running time then is $O(n^2 2^{2n})$.
As before the evaluation cost of $a(X)$ can be reduced by going through the 2^{2n} binary strings in Gray code order. The advantage using Larsen's method over Godfrey's is that there is no longer an addition factor 2^{2n}. However Larsen's experiments show that Godfrey's procedure is faster in practice.

3 Results

We implement Godfrey's procedure for computing the number solutions to the Langford and the Nickerson's variant on NVIDIA's CUDA parallel computing platform. The calcualtion is done modulo primes then the actual value is obtained by chinese remainder theorem. We checked or implementation by computing the previously known values. New results are found for the original problem when $n = 27$, $n = 28$. We note that recently the 27^{th} langford number was reported, and does not match with our results. It is worth noting that our results match in the order of magituide and the seven most significant digits. We also found new results for the variant when $n = 24$, $n = 25$, and partially computed the value for $n = 28$. These correspond to sequences $A014552$ and $A059106$ in Sloane's *Online Encyclopedia of Integer Sequences*. Additionally the corresponding values for the sequence of number $A176127$ and $A004075$ in the *oeis* are just twice that of $A014552$ and $A059106$ respectively, since in the orignal problem refelected solutions were considered the same. The table below shows the sequences along with our results in red.

4 Acknowledgements

We like to thank the CUNY HPCC for the use of their computational cluster. The CUNY HPCC is operated by the College of Staten Island and funded, in part, by grants from the City of New York, State of New York, CUNY Research Foundation, and National Science Foundation Grants CNS-0958379, CNS-0855217 and ACI 1126113.

References

[1] http://www.dialectrix.com/langford.html

[2] C. D. Langford. Problem. The Mathematical Gazette, 42,1958, 228

[3] T. Skolem, On Certain Distributions of Integers in Pairs with Given Differences, Math. Scand. 5, 1957, 57-68

[4] C. J. Priday, Langford's Problem. I., Math. Gaz., 1959, v43, 250-3

[5] R. O. Davies, Langford's problem. II., Math. Gaz., 1959, v43, 253-5

[6] R.S Nickerson, Problem e1845, Amer. Math. Monthly, 73, 1966, 81

[7] N. J. A. Sloane, The online encyclopedia of integer sequences, http://oeis.org/

[8] J. W. Larsen, Counting the number of skolem sequences using inclusion-exclusion, 2009

[9] Z. Habbas, M. Krajecki, and D. Singer. The langfords problem: A challenge for parallel resolution of csp. In Parallel Processing and Applied Mathematics, pages 789–796. Springer, 2002