TERNARY WORDS AND JACOBSTHAL NUMBERS

THOMAS KOSHY AND RALPH P. GRIMALDI

Abstract. We investigate a special class of ternary words, and explore some close and interesting relationships between them and the well-known Jacobsthal numbers.

1. Introduction

1.1. **Jacobsthal Numbers.** The *Jacobsthal numbers*, named after the German mathematician Ernst Erich Jacobsthal (1882–1965), and the *Jacobsthal-Lucas numbers* satisfy the recurrence $x_n = x_{n-1} + 2x_{n-2}$, where $n \geq 3$. When $x_1 = 1 = x_2, x_n = J_n$, the nth Jacobsthal number; when $x_1 = 1$ and $x_2 = 5, x_n = j_n$, the nth Jacobsthal-Lucas number. It follows by the Jacobsthal recurrence that $J_0 = 0, J_{-1} = 1/2, j_0 = 2$, and $j_{-1} = -1/2$.

Both J_n and j_n can also be defined explicitly by the *Binet-like* formulas $J_n = \dfrac{2^n - (-1)^n}{3}$, and $j_n = 2^n + (-1)^n$, where n is any integer. Table 1 shows twelve Jacobsthal and Jacobsthal-Lucas numbers, where $-1 \leq n \leq 10$.

Table 1: Jacobsthal and Jacobsthal-Lucas Numbers

n	-1	0	1	2	3	4	5	6	7	8	9	10
J_n	1/2	0	1	1	3	5	11	21	43	85	171	341
j_n	$-1/2$	2	1	5	7	17	31	65	127	257	511	1025

Using the Binet-like formulas and Jacobsthal recurrence, we can extract an array of interesting properties [5]. For example, $J_n + J_{n+1} = 2^n, J_{n+1} - 2J_n = (-1)^n$, and $J_{n+1} + 2J_{n-1} = j_n$.

1.2. **Formal Languages.** An *alphabet* Σ is a finite set of symbols. A *word* (or *string*) *over* Σ is a finite sequence of symbols from Σ. The number of symbols in a word is its *length*. The word of length 0 is the *empty word* or *null word*; it is denoted by λ.

The set of all possible words over Σ, denoted by Σ^*, is the *Kleene closure* of Σ; it is named after the American logician Stephen Kleene (1909–1994). A *language* L over Σ is a subset of Σ^*.

The *concatenation* of two words x and y in L, denoted by xy, is obtained by appending y at the end of x. For example, the concatenation of $x = x_1x_2 \ldots x_m$ and $y = y_1y_2 \ldots y_n$ is $xy = x_1x_2 \ldots x_my_1y_2 \ldots y_n$. The *concatenation* of two languages A and B over Σ, denoted by AB, is defined by $AB = \{ab|a \in A \text{ and } b \in B\}$. In particular, $A^2 = \{ab|a, b \in A\}$. More generally, $A^n = \{a_1a_2 \ldots a_n|a_i \in A, 1 \leq i \leq n\}$ and $A^0 = \{\lambda\}$. Then $A^* = \bigcup_{n=0}^{\infty} A^n$.

In particular, let $\Sigma = \{0, 1\}$, the *binary alphabet*; its symbols are the *bits* 0 and 1. Let $L = \{0, 01, 11\}$. There are exactly J_{n+1} words of length n in L^*, where $n \geq 1$ [3].

2. A Ternary Version

We now pursue a ternary version of the binary case, but with some added restrictions. It appeared in the final round of the 1987 Austrian Olympiad [1, 4]. It is interesting in its own right and has fascinating implications.

Let $\Sigma = \{0, 1, 2\}$. The digits 0, 1, and 2 are *ternary digits*. (In the Austrian Olympiad problem, $\Sigma = \{a, b, c\}$.) Let b_n denote the number of *ternary words* $w_n = x_1x_2 \dots x_n$ of length n such that $x_1 = 0 = x_n$ and $x_i \neq x_{i+1}$, where $x_i \in \Sigma$ and $1 \leq i \leq n-1$. Clearly, the *reverse* w_n^R of an acceptable word $w_n = 0x_2 \dots x_{n-1}0$ is also acceptable. (*Note*: In the interest of brevity and convenience, in the rest of the article, "ternary words" will mean "ternary words with the added restrictions," when there is *no* ambiguity.)

Table 2 lists the ternary words w_n and the corresponding numbers b_n, where $1 \leq n \leq 6$. Notice that there are *no* ternary words of length 2 that satisfy the given conditions. Although the counts b_n do not seem to follow a pattern, the following theorem establishes a simple formula for b_n using a constructive algorithm.

Table 2: Ternary Words and Their Counts

n	Ternary Words w_n	b_n
1	0	1
2	.	0
3	010, 020	2
4	0120, 0210	2
5	01210, 02120	6
	01010, 02010	
	01020, 02020	
6	010120, 010210, 020120, 020210, 012120, 021210	10
	012010, 021010	
	012020, 021020	

Theorem 2.1. *Let b_n denote the number of ternary words $w_n = x_1x_2 \dots x_n$ of length n such that $x_1 = 0 = x_n$ and $x_i \neq x_{i+1}$, where $1 \leq i \leq n-1$. Then $b_n = 2J_{n-2}$, where $n \geq 1$.*

Proof. It is easy to confirm the claim for $1 \leq n \leq 4$. Let w_n be an arbitrary ternary word of length $n \geq 5$. We will now employ an algorithm to construct words of length n from those of lengths $n-1$ and $n-2$.

Step 1. Replace the last digit $x_{n-1} = 0$ in w_{n-1} with 10 if $x_{n-2} = 2$; otherwise, replace it with 20.
Step 2A. Append 10 at the end of each w_{n-2}.
Step 2B. Append 20 at the end of each w_{n-2}.

Since the algorithm is reversible, it produces all desired ternary words w_n.

Step 1 yields b_{n-1} words w_n. Steps 2A and 2B produce b_{n-2} words each. Thus, $b_n = b_{n-1} + 2b_{n-2}$. This recurrence, paired with the initial conditions, gives the desired result. □

We will now illustrate the steps in the proof for the case $n = 6$.
Step 1. There are three words $w_5 = 0x_2x_3x_40$ with $x_4 = 2$; replace each $x_5 = 0$ with 10. The three remaining words have $x_4 = 1$; replace each $x_5 = 0$ with 20:

02120	01020	02020	01210	01010	02010
↓	↓	↓	↓	↓	↓
0212**10**	0102**10**	0202**10**	0121**20**	0101**20**	0201**20**.

Step 2A. Append 10 at the end of each w_4:

$$\begin{array}{cc} 0120 & 0210 \\ \downarrow & \downarrow \\ 0120\mathbf{10} & 0210\mathbf{10}. \end{array}$$

Step 2B. Append 20 at the end of each w_4:

$$\begin{array}{cc} 0120 & 0210 \\ \downarrow & \downarrow \\ 0120\mathbf{20} & 0210\mathbf{20}. \end{array}$$

Clearly, these steps produce the $b_6 = 10$ ternary words.

The following result is an immediate consequence of the constructive algorithm.

Corollary 1. *There are exactly $\frac{1}{2}b_n = J_{n-2}$ ternary words w_n that begin with 01 (or end in 10), where $n \geq 2$.* □

The next result follows from this corollary and we will use it several times in our discourse.

Corollary 2. *There are J_{n-2} ternary words w_n that begin with 02 (or end in 20), where $n \geq 2$.* □

We now have the needed machinery to develop an explicit formula for the number of 0's among the b_n ternary words of length n.

2.1. **Zeros Among the b_n Ternary Numbers.** Let z_n denote the number of 0's among the b_n ternary words w_n of length n. For example, $z_1 = 1, z_2 = 0, z_3 = 4 = z_4, z_5 = 16$, and $z_6 = 28$; see Table 2.

Using the above constructive algorithm, we can easily develop a recurrence for z_n. Replacing x_{n-1} in w_{n-1} in Step 1 with 10 or 20 does not contribute any new 0's. So Step 1 contributes z_{n-1} 0s to z_n. Each of Steps 2A and 2B contributes $z_{n-2} + b_{n-2}$ zeros to z_n. Thus,

$$\begin{aligned} z_n &= z_{n-1} + 2(z_{n-2} + b_{n-2}) \\ &= z_{n-1} + 2z_{n-2} + 4J_{n-4} \\ &= z_{n-1} + 2z_{n-2} + \frac{4}{3}\left[2^{n-4} - (-1)^{n-4}\right], \end{aligned} \tag{2.1}$$

where $z_1 = 1, z_2 = 0$, and $n \geq 3$.

The general solution of recurrence (2.1) is of the form $z_n = c_1 \cdot 2^n + c_2(-1)^n + An2^n + Bn(-1)^n$ [2, 6]. Substituting $An2^n$ in the recurrence $z_n = z_{n-1} + 2z_{n-2} + \left(\frac{4}{3}\right)2^{n-4}$ yields $A = 1/18$. Likewise, substituting $Bn(-1)^n$ in the recurrence $z_n = z_{n-1} + 2z_{n-2} - \frac{4}{3}(-1)^{n-4}$ yields $B = -4/9$. Thus,

$$z_n = c_1 \cdot 2^n + c_2(-1)^n + \left(\frac{n}{18}\right)2^n - \left(\frac{4n}{9}\right)(-1)^n.$$

Using the initial conditions $z_1 = 1$ and $z_2 = 0$, this yields $c_1 = \dfrac{8}{54} = -c_2$. Thus,

$$\begin{aligned} z_n &= \left(\frac{8}{54}\right)2^n - \left(\frac{8}{54}\right)(-1)^n + \left(\frac{n}{18}\right)2^n - \left(\frac{4n}{9}\right)(-1)^n \\ &= \left(\frac{3n+8}{54}\right)2^n - \left(\frac{24n+8}{54}\right)(-1)^n \\ &= \left(\frac{3n+8}{54}\right)(J_n + J_{n+1}) - \left(\frac{24n+8}{54}\right)(J_{n+1} - 2J_n) \\ &= \left(\frac{17n+8}{18}\right)J_n - \left(\frac{7n}{18}\right)J_{n+1}, \end{aligned} \tag{2.2}$$

where $n \geq 1$.

For example, $z_{10} = \dfrac{178 \cdot 341}{18} - \dfrac{70 \cdot 683}{18} = 716$.

2.2. **Nonzero Digits Among the b_n Ternary Numbers.** It follows from formula (2.2) that the number of nonzero digits $nonz_n$ among the b_n ternary numbers w_n is given by

$$\begin{aligned} nonz_n &= nb_n - z_n \\ &= n(2J_{n-2}) - \left[\left(\frac{17n+8}{18}\right)J_n - \left(\frac{7n}{18}\right)J_{n+1}\right] \\ &= \left(\frac{7n}{18}\right)J_{n+1} - \left(\frac{17n+8}{18}\right)J_n + 2nJ_{n-2}. \end{aligned} \tag{2.3}$$

Consequently,

$$\begin{aligned} \text{the number of 1's} &= \text{the number of 2's} \\ &= \frac{1}{2}(nb_n - z_n) \\ &= \left(\frac{7n}{36}\right)J_{n+1} - \left(\frac{17n+8}{36}\right)J_n + nJ_{n-2}. \end{aligned}$$

For example, $nonz_6 = \left(\dfrac{7 \cdot 6}{18}\right)J_7 - \left(\dfrac{17 \cdot 6 + 8}{18}\right)J_6 + 12J_4 = \dfrac{42 \cdot 43}{18} - \dfrac{110 \cdot 21}{18} + 12 \cdot 5 = 32$, as found in Table 2. Further, there are 16 1's and 16 2's.

Next, we compute the cumulative sum of the decimal values of the b_n ternary words when considered as ternary numbers. We will accomplish this using recursion and the constructive algorithm.

2.3. **Cumulative Sum of the b_n Ternary Numbers.** Let S_n denote the cumulative sum of the decimal values of the b_n ternary numbers. It follows from Table 2 that $S_1 = 0 = S_2, S_3 = 9, S_4 = 36$, and $S_5 = 297$. Let $w_k = 0x_2x_3 \ldots x_{k-1}0$ be an arbitrary ternary number with k digits.

Step 1. Replacing x_{n-1} with 10 or 20 shifts $0x_2 \ldots x_{n-2}$ two places to the left of 10, or 20, respectively. Since there are b_{n-1} ternary numbers with $n-1$ digits, this step contributes $3S_{n-1} + 2 \cdot 3\left(\frac{1}{2}b_{n-1}\right) + 1 \cdot 3\left(\frac{1}{2}b_{n-1}\right) = 3S_{n-1} + \left(\frac{9}{2}\right)b_{n-1} = 3S_{n-1} + 9J_{n-3}$ to the sum S_n.

Step 2A. Appending 10 at the end of $0x_2 \ldots x_{n-3}0$ shifts it two positions to the left. The contribution resulting from this operation is $3^2S_{n-2} + 1 \cdot 3b_{n-2} = 9S_{n-2} + 6J_{n-4}$.

Step 2B. Appending 20 at the end of $0x_2 \ldots x_{n-3}0$ contributes $3^2S_{n-2} + 2 \cdot 3b_{n-2} = 9S_{n-2} +$

$12J_{n-4}$ to the grand total.

Combining these steps, we get

$$\begin{aligned} S_n &= (3S_{n-1} + 9J_{n-3}) + (9S_{n-2} + 6J_{n-4}) + (9S_{n-2} + 12J_{n-4}) \\ &= 3S_{n-1} + 18S_{n-2} + 9J_{n-3} + 18J_{n-4} \\ &= 3S_{n-1} + 18S_{n-2} + 9J_{n-2}, \end{aligned} \tag{2.4}$$

where $n \geq 4$.

For example, $S_5 = 3S_4 + 18S_3 + 9J_3 = 3 \cdot 36 + 18 \cdot 9 + 9 \cdot 3 = 297$.

2.4. **An Explicit Formula for S_n.** It follows from recurrence (2.4) that

$$S_n = 3S_{n-1} + 18S_{n-2} + 3 \cdot 2^{n-2} - 3(-1)^n. \tag{2.5}$$

The roots of the characteristic equation of the homogeneous recurrence $S_n = 3S_{n-1} + 18S_{n-2}$ are -3 and 6. The particular part of the solution of recurrence (2.5) corresponding to the nonhomogeneous part $3 \cdot 2^{n-2}$ has the form $A \cdot 2^n$. Substituting this in the recurrence $S_n = 3S_{n-1}+18S_{n-2}+3\cdot 2^{n-2}$ yields $A = -3/20$. The particular part of the solution of this recurrence corresponding to the nonhomogeneous part $-3(-1)^n$ has the form $B(-1)^n$. Substituting this in the recurrence $S_n = 3S_{n-1} + 18S_{n-2} - 3(-1)^n$ yields $B = 3/14$.

The general solution of recurrence (2.5) is of the form

$$S_n = c_1(-3)^n + c_2 \cdot 6^n - \left(\frac{3}{20}\right) 2^n + \left(\frac{3}{14}\right)(-1)^n.$$

Using the initial conditions $S_1 = 0 = S_2$, this recurrence yields $c_1 = -\dfrac{14}{140}$ and $c_2 = \dfrac{5}{140}$. Thus,

$$S_n = -\left(\frac{14}{140}\right)(-3)^n + \left(\frac{5}{140}\right)6^n - \left(\frac{3}{20}\right)2^n + \left(\frac{3}{14}\right)(-1)^n, \tag{2.6}$$

where $n \geq 1$.

For example, $S_5 = \dfrac{14 \cdot 3^5 + 5 \cdot 6^5}{140} - \dfrac{48 \cdot 14 + 3 \cdot 10}{140} = 297$, as expected.

Since $J_n + J_{n+1} = 2^n$ and $J_{n+1} - 2J_n = (-1)^n$, formula (2.6) can be rewritten in terms of Jacobsthal numbers. For convenience, we now let $a = -14/140, b = 5/140, c = -3/20$, and $d = 3/14$. Then

$$\begin{aligned} S_n &= (3^n b + c)(J_n + J_{n+1}) + (3^n a + d)(J_{n+1} - 2J_n) \\ &= [3^n(a+b) + c + d]J_{n+1} + [3^n(b - 2a) + c - 2d]J_n \\ &= \frac{1}{140}\left[(-3)^{n+2} + 9\right] J_{n+1} - \frac{1}{140}\left[3^n(-33) + 81\right] J_n \\ &= \frac{1}{140}\left(11 \cdot 3^{n+1} - 81\right) J_n - \frac{1}{140}\left(3^{n+2} - 9\right) J_{n+1}. \end{aligned} \tag{2.7}$$

For example, $S_4 = \dfrac{1}{140}\left(11 \cdot 3^5 - 81\right) \cdot 5 - \dfrac{1}{140}\left(3^6 - 9\right) \cdot 11 = \dfrac{12,960 - 7,920}{140} = 36$, again as expected.

3. Inversions

Next we investigate the number of inversions in words over Σ. To begin with, let $\{a_1, a_2, \ldots, a_n\}$ be a *totally ordered* alphabet with $a_1 < a_2 < \cdots < a_n$. Let $x_1x_2\cdots x_k$ be a word of length k over this alphabet. For $1 \le i < j \le k$, call the pair x_i and x_j an *inversion* if $x_i > x_j$.

For example, let $\Sigma = \{0, 1, 2\}$, where $0 < 1 < 2$. Then the word $x_1x_2x_3x_4x_5 = 01210$ contains four inversions: $x_2 > x_5, x_3 > x_4, x_3 > x_5$, and $x_4 > x_5$.

Let inv_n count the number of inversions among the b_n ternary words of length n, where $n \ge 3$. Then $inv_3 = 2, inv_4 = 5, inv_5 = 21, inv_6 = 56$, and $inv_7 = 164$. We will now establish that inv_n satisfies the recurrence

$$inv_n = inv_{n-1} + b_{n-1} + \frac{1}{2}(nonz_{n-1} - nonz_{n-2}) + 2inv_{n-2} + 2b_{n-2} + 2nonz_{n-2} + \frac{1}{2}nonz_{n-2},$$

where $n \ge 3$.

1) From Step 1 of the algorithm when the last digit x_{n-1} is replaced (by either 10 or 20), a new inversion arises. These are counted by b_{n-1}.
2) Also from Step 1 of the algorithm, for the case where the last digit x_{n-1} is replaced by 10 (for when $x_{n-2} = 2$), there is a new inversion for each 2 that occurs among the first $n-2$ digits of the b_{n-1} words that end in 20. These inversions are counted by $\frac{1}{2}(nonz_{n-1} - nonz_{n-2})$.
3) Steps 2A and 2B of the algorithm each provide b_{n-2} new inversions: 10 in positions $n-1$ and n for Step 2A; 20 in positions $n-1$ and n for Step 2B.
4) Each of Steps 2A and 2B of the algorithm provides $nonz_{n-2}$ new inversions with the new 0 now in position n.
5) Finally, from Step 2A, we get $\frac{1}{2}nonz_{n-2}$ new inversions for each of the $\frac{1}{2}nonz_{n-2}$ 2s that occur among the first $n-2$ positions of the b_n words that end in 10. Each such 2 provides an inversion with the new 1 in position $n-1$.

Combining these five steps, we get

$$\begin{aligned}
inv_n &= inv_{n-1} + 2inv_{n-2} + b_{n-1} + 2b_{n-2} + \frac{1}{2}nonz_{n-1} + 2nonz_{n-2}\\
&= inv_{n-1} + 2inv_{n-2} + 2J_{n-3} + 4J_{n-4}\\
&\quad + \frac{1}{2}\left\{(n-1)(2J_{n-3}) - \left[\left(\frac{17(n-1)+8}{18}\right)J_{n-1} - \left(\frac{7(n-1)}{18}\right)J_n\right]\right\}\\
&\quad + 2\left\{(n-2)(2J_{n-4}) - \left[\left(\frac{17(n-2)+8}{18}\right)J_{n-2} - \left(\frac{7(n-2)}{18}\right)J_{n-1}\right]\right\}.
\end{aligned}$$

Substituting for J_n, yields

$$inv_n = inv_{n-1} + 2inv_{n-2} + \frac{1}{3}n(-1)^{n+1} + \frac{1}{12}n(2^n) + \frac{1}{3}(-1)^n - \frac{1}{12}(2^n).$$

Consequently, the general solution of the recurrence is of the form

$$inv_n = c_1(2^n) + c_2(-1)^n + An2^n + Bn^22^n + Cn(-1)^n + Dn^2(-1)^n.$$

Next we will determine the coefficients for the particular part of the solution.

1) To find A and B, substitute $inv_n = An2^n + Bn^22^n$ in the recurrence $inv_n = inv_{n-1} +$

$2inv_{n-2} + \frac{1}{12}n(2^n) - \frac{1}{12}(2^n)$. After some basic algebra, this gives

$$0 = -\frac{3}{2}A(2^n) - 3Bn(2^n) + \frac{5}{2}B(2^n) + \frac{1}{12}n(2^n) - \frac{1}{12}(2^n).$$

Comparing the coefficients for 2^n and $n2^n$, we get $0 = -\frac{3}{2}A + \frac{5}{2}B - \frac{1}{12}$ and $0 = -3B + \frac{1}{12}$, so $A = -\frac{1}{108}$ and $B = \frac{1}{36}$.

2) To find C and D, substitute $inv_n = Cn(-1)^n + Dn^2(-1)^n$ in the recurrence $inv_n = inv_{n-1} + 2inv_{n-2} + \frac{1}{3}n(-1)^{n+1} + \frac{1}{3}(-1)^n$. After some simplification, this yields

$$0 = -3C(-1)^n - 6Dn(-1)^n + 7D(-1)^n - \frac{1}{3}n(-1)^n + \frac{1}{3}(-1)^n.$$

Comparing the coefficients for $(-1)^n$ and $n(-1)^n$, we get $0 = -3C + 7D + \frac{1}{3}$ and $0 = -6D - \frac{1}{3}$, so $C = -\frac{1}{54}$ and $D = -\frac{1}{18}$.

Consequently,

$$inv_n = c_1(2^n) + c_2(-1)^n + \left(-\frac{1}{108}\right)n2^n + \left(\frac{1}{36}\right)n^2 2^n + \left(-\frac{1}{54}\right)n(-1)^n + \left(-\frac{1}{18}\right)n^2(-1)^n.$$

3) The initial conditions $inv_3 = 2$ and $inv_4 = 5$ yield $c_1 = -\frac{1}{27} = -c_2$. Thus,

$$\begin{aligned} inv_n &= -\frac{1}{27}(2^n) + \frac{1}{27}(-1)^n + \left(-\frac{1}{108}\right)n2^n + \left(\frac{1}{36}\right)n^2 2^n \\ &\quad + \left(-\frac{1}{54}\right)n(-1)^n + \left(-\frac{1}{18}\right)n^2(-1)^n \\ &= -\frac{1}{9}J_n + \frac{n}{36}\left(J_n - J_{n+1}\right) + \frac{n^2}{36}\left(5J_n - J_{n+1}\right) \\ &= \left(\frac{5n^2 + n - 4}{36}\right)J_n - \left(\frac{n^2 + n}{36}\right)J_{n+1}, \end{aligned}$$

where $n \geq 3$.

4. Acknowledgment

The authors would like to thank the referee for his/her carefully worded comments and suggestions for improving the quality of exposition of the original version.

References

[1] *1987 Austrian Olympiad*, Crux Mathematicorum, **15** (1989), 264.
[2] R. P. Grimaldi, *Discrete and Combinatorial Mathematics*, 5th edition, Pearson, Boston, Massachusetts, 2004.
[3] R. P. Grimaldi, *Binary strings and the Jacobsthal numbers*, Congressus Numerantium, **174** (2005), 3–22.
[4] R. Honsberger, *From Erdös to Kiev*, Mathematical Association of America, Washington, D. C., 1996.
[5] A. F. Horadam, *Jacobsthal representation numbers*, The Fibonacci Quarterly, **34.1** (1996), 40–54.
[6] T. Koshy, *Discrete Mathematics with Applications*, Elsevier, Boston, Massachusetts, 2004.

MSC2010: 11B37, 11B39, 11B50, 39A06

Department of Mathematics, Framingham State University, Framingham, MA 01701
E-mail address: tkoshy@emeriti.framingham.edu

Department of Mathematics, Rose-Hulman Institute of Technology, Terre Haute, IN 47808
E-mail address: grimaldi@rose-hulman.edu