

# Thue-Morse Sequence

Sava Darius-Ștefan

Faculty of Computer Science  
University of Bucharest

27 May, 2025

# Table of Contents

- 1 How to build this sequence
- 2 Applications
- 3 Combinatorial Properties
- 4 Math Problem
- 5 Fair Share Sequence
- 6 Magic Squares
- 7 Pattern Avoidance
- 8 Music
- 9 An interesting problem

On Monday the baby said A, on Tuesday AU, on Wednesday AUUA, on Thursday AUUAUAAU. What will he say on Friday?

You can see that this very gifted baby increases her talking capacity twice each day. If the baby continues indefinitely, her text converges to an infinite sequence that mathematicians call the Thue-Morse sequence. Of course, mathematicians use zeroes and ones instead of A and U, so the sequence looks like 0110100110010110100 ...

# Sequence construction

The Thue-Morse sequence, denoted as  $(t_n)_{n=0}^{\infty}$ , can be constructed using the following methods:

① **Append the complement of the previous sequence:**

Start with 0, and iteratively append the complement of the sequence generated so far:

0, 01, 0110, 01101001, 0110100110010110, ...

## ② Lindenmayer system (L-system)

Here is an alternate way to construct the sequence. In the same manner, we can start with 01 and then we continually walk through the sequence, taking the next entry that we haven't visited yet (we consider that the first 0 is visited), placing two copies at the end, and flipping the second bit.

01

$0 \rightarrow 00 \rightarrow 01$

$1 \rightarrow 11 \rightarrow 10$

## ② Lindenmayer system (L-system)

Here is an alternate way to construct the sequence. In the same manner, we can start with 01 and then we continually walk through the sequence, taking the next entry that we haven't visited yet (we consider that the first 0 is visited), placing two copies at the end, and flipping the second bit.

0110

$0 \rightarrow 00 \rightarrow 01$

$1 \rightarrow 11 \rightarrow 10$

## ② Lindenmayer system (L-system)

Here is an alternate way to construct the sequence. In the same manner, we can start with 01 and then we continually walk through the sequence, taking the next entry that we haven't visited yet (we consider that the first 0 is visited), placing two copies at the end, and flipping the second bit.

011010

$0 \rightarrow 00 \rightarrow 01$

$1 \rightarrow 11 \rightarrow 10$

## ② Lindenmayer system (L-system)

Here is an alternate way to construct the sequence. In the same manner, we can start with 01 and then we continually walk through the sequence, taking the next entry that we haven't visited yet (we consider that the first 0 is visited), placing two copies at the end, and flipping the second bit.

01101001

0  $\rightarrow$  00  $\rightarrow$  01

1  $\rightarrow$  11  $\rightarrow$  10



## ② Lindenmayer system (L-system)

Here is an alternate way to construct the sequence. In the same manner, we can start with 01 and then we continually walk through the sequence, taking the next entry that we haven't visited yet (we consider that the first 0 is visited), placing two copies at the end, and flipping the second bit.

0110100110

0  $\rightarrow$  00  $\rightarrow$  01

1  $\rightarrow$  11  $\rightarrow$  10

## ② Lindenmayer system (L-system)

Here is an alternate way to construct the sequence. In the same manner, we can start with 01 and then we continually walk through the sequence, taking the next entry that we haven't visited yet (we consider that the first 0 is visited), placing two copies at the end, and flipping the second bit.

011010011001

$0 \rightarrow 00 \rightarrow 01$

$1 \rightarrow 11 \rightarrow 10$

# Sequence Construction

If we do this indefinitely, we once again will end up with the Thue-Morse Sequence. As we can see, this is matching what we generated in the previous slide. This construction shows us that the  $2n$ -th entry of the Thue-Morse Sequence is equal to the  $n$ -th entry, and the  $2n + 1$ -th entry of the Thue-Morse Sequence is equal to  $1 -$  the  $n$ -th entry, starting with  $t_0 = 0$ .

We now have a recursive formula that we will make use of later.

$$t_{2n} = t_n; \quad t_{2n+1} = 1 - t_n; \quad \text{and} \quad t_0 = 0.$$

## 3 Using the binary representation

The third technique to generate the Thue-Morse Sequence comes from computing binary representations. We start with a number, we represent it in binary, we take the base two digit sum and then decide if that's even or odd.

0

Number	Binary Representation	Base 2 Digital Sum	Parity of Digital Sum
0	0	0	0

## 3 Using the binary representation

The third technique to generate the Thue-Morse Sequence comes from computing binary representations. We start with a number, we represent it in binary, we take the base two digit sum and then decide if that's even or odd.

01

Number	Binary Representation	Base 2 Digital Sum	Parity of Digital Sum
1	1	1	1

## 3 Using the binary representation

The third technique to generate the Thue-Morse Sequence comes from computing binary representations. We start with a number, we represent it in binary, we take the base two digit sum and then decide if that's even or odd.

011

Number	Binary Representation	Base 2 Digital Sum	Parity of Digital Sum
2	10	1	1

## 3 Using the binary representation

The third technique to generate the Thue-Morse Sequence comes from computing binary representations. We start with a number, we represent it in binary, we take the base two digit sum and then decide if that's even or odd.

0110

Number	Binary Representation	Base 2 Digital Sum	Parity of Digital Sum
3	11	2	0

## 3 Using the binary representation

The third technique to generate the Thue-Morse Sequence comes from computing binary representations. We start with a number, we represent it in binary, we take the base two digit sum and then decide if that's even or odd.

01101

Number	Binary Representation	Base 2 Digital Sum	Parity of Digital Sum
4	100	1	1



## 3 Using the binary representation

The third technique to generate the Thue-Morse Sequence comes from computing binary representations. We start with a number, we represent it in binary, we take the base two digit sum and then decide if that's even or odd.

011010

Number	Binary Representation	Base 2 Digital Sum	Parity of Digital Sum
5	101	2	0

## 3 Using the binary representation

The third technique to generate the Thue-Morse Sequence comes from computing binary representations. We start with a number, we represent it in binary, we take the base two digit sum and then decide if that's even or odd.

0110100

Number	Binary Representation	Base 2 Digital Sum	Parity of Digital Sum
6	110	2	0

## 3 Using the binary representation

The third technique to generate the Thue-Morse Sequence comes from computing binary representations. We start with a number, we represent it in binary, we take the base two digit sum and then decide if that's even or odd.

01101001 ...

Number	Binary Representation	Base 2 Digital Sum	Parity of Digital Sum
7	111	3	1

$$t_n = s_2(n) \pmod{2}$$

# Applications

The Thue–Morse sequence has found applications in a variety of domains:

- **Mathematics:** Serves as a counterexample in several problems and helps study the structure of certain algebraic objects.
- **Computer Science:**
  - *Pattern Avoidance:* Helps design strings that avoid repeated patterns in text algorithms.
  - *Automata Theory:* A classic example of an automatic sequence, generated by finite automata.
  - *Cryptography:* Due to its balance and lack of repetition, the sequence helps in generating pseudo-random key streams and designing cryptographic systems that resist pattern-based attacks.
  - *Data Structures:* Supports balanced and non-repetitive data layouts.
- **Music and Art:** The sequence inspires compositions and artworks that explore symmetry, pattern. Used to create compositions and rhythms that are structured yet non-repetitive.

The Thue–Morse Sequence's simple definition contrasts with its deep and far-reaching theoretical and practical implications.

# Combinatorial Properties

Consider the integers 0 up to  $2^n - 1$ .

As shown below, we consider the integers from 0 to 15 (which is  $2^4 - 1$ ).

Below these integers, we place the Thue-Morse Sequence:

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	1	1	0	1	0	0	1	1	0	0	1	0	1	1	0

Now let's divide them into two groups:

- **Group 1 (Numbers with 1 in the sequence):**

1, 2, 4, 7, 8, 11, 13, 14

- **Group 2 (Numbers with 0 in the sequence):**

0, 3, 5, 6, 9, 10, 12, 15

1 2 4 7 8 11 13 14

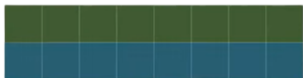
0 3 5 6 9 10 12 15

1 2 4 7 8 11 13 14



0 3 5 6 9 10 12 15

1 2 4 7 8 11 13 14



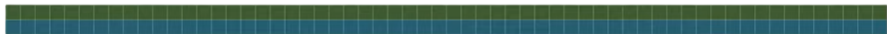
0 3 5 6 9 10 12 15



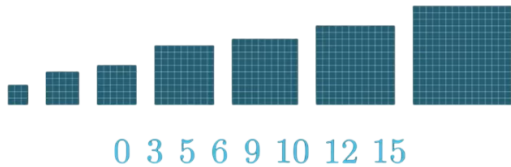
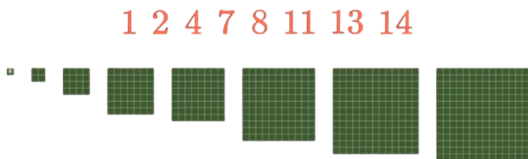
1 2 4 7 8 11 13 14

0 3 5 6 9 10 12 15

1 2 4 7 8 11 13 14



0 3 5 6 9 10 12 15



1 2 4 7 8 11 13 14



0 3 5 6 9 10 12 15

**1      2      4      7      8      10      13      14**

$$0^0 + 3^0 + 5^0 + 6^0 + 9^0 + 10^0 + 12^0 + 15^0 = 1^0 + 2^0 + 4^0 + 7^0 + 8^0 + 11^0 + 13^0 + 14^0$$

$$0^1 + 3^1 + 5^1 + 6^1 + 9^1 + 10^1 + 12^1 + 15^1 = 1^1 + 2^1 + 4^1 + 7^1 + 8^1 + 11^1 + 13^1 + 14^1$$

$$0^2 + 3^2 + 5^2 + 6^2 + 9^2 + 10^2 + 12^2 + 15^2 = 1^2 + 2^2 + 4^2 + 7^2 + 8^2 + 11^2 + 13^2 + 14^2$$

$$0^3 + 3^3 + 5^3 + 6^3 + 9^3 + 10^3 + 12^3 + 15^3 = 1^3 + 2^3 + 4^3 + 7^3 + 8^3 + 11^3 + 13^3 + 14^3$$

**0      3      5      6      9      11      12      15**

## Theorem

*The Thue-Morse sequence  $t = (t_n)_{n \geq 0}$  has the following property.  
Define*

$$I = \{0 \leq i < 2^N \mid t_i = 0\}$$

$$J = \{0 \leq j < 2^N \mid t_j = 1\}$$

*Then for  $0 \leq k < N$  we have*

$$\sum_{i \in I} i^k = \sum_{j \in J} j^k.$$

There are many other interesting occurrences of this sequence but let's see one in mathematics.

$$\left(\frac{1}{2}\right) \left(\frac{3}{4}\right) \left(\frac{5}{6}\right) \left(\frac{7}{8}\right) \left(\frac{9}{10}\right) \left(\frac{11}{12}\right) \left(\frac{13}{14}\right) \cdots \rightarrow 0$$

$$\left(\frac{2}{1}\right) \left(\frac{4}{3}\right) \left(\frac{6}{5}\right) \left(\frac{8}{7}\right) \left(\frac{10}{9}\right) \left(\frac{12}{11}\right) \left(\frac{14}{13}\right) \cdots \rightarrow \infty$$

# Math Problem

$$\begin{array}{cccccccc} \left(\frac{1}{2}\right) & \left(\frac{3}{4}\right) & \left(\frac{5}{6}\right) & \left(\frac{7}{8}\right) & \left(\frac{9}{10}\right) & \left(\frac{11}{12}\right) & \left(\frac{13}{14}\right) & \left(\frac{15}{16}\right) \cdots \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{array}$$

$$\begin{array}{cccccccc} \left(\frac{1}{2}\right) & \left(\frac{4}{3}\right) & \left(\frac{6}{5}\right) & \left(\frac{7}{8}\right) & \left(\frac{10}{9}\right) & \left(\frac{11}{12}\right) & \left(\frac{13}{14}\right) & \left(\frac{16}{15}\right) \cdots \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{array}$$



## Theorem

If  $\varepsilon_n = (-1)^{t_n}$ , then:

$$\prod_{n=0}^{\infty} \left( \frac{2n+1}{2n+2} \right)^{\varepsilon_n} = \frac{\sqrt{2}}{2}$$

**Proof.** Define: P and Q as follows:

$$P = \prod_{n=0}^{\infty} \left( \frac{2n+1}{2n+2} \right)^{\varepsilon_n}, \quad Q = \prod_{n=1}^{\infty} \left( \frac{2n}{2n+1} \right)^{\varepsilon_n}$$

Now, consider the product  $PQ$ :

$$PQ = \prod_{n=0}^{\infty} \left( \frac{2n+1}{2n+2} \right)^{\varepsilon_n} \cdot \prod_{n=1}^{\infty} \left( \frac{2n}{2n+1} \right)^{\varepsilon_n}$$

# Math Problem

We can replace this with the infinite product from  $n = 1$  to  $\infty$  of the two expressions  $P$  and  $Q$  multiplied together, where we leave the  $\frac{1}{2}$  out in front from the  $n = 0$  term of the product for  $P$ .

$$PQ = \frac{1}{2} \prod_{n=1}^{\infty} \left( \frac{2n+1}{2n+2} \cdot \frac{2n}{2n+1} \right)^{\epsilon_n}.$$

$$PQ = \frac{1}{2} \prod_{n=1}^{\infty} \left( \frac{n}{n+1} \right)^{\epsilon_n}$$

Now we break this infinite product into separate products over odd and even indices, we find :

$$PQ = \frac{1}{2} \prod_{n=1}^{\infty} \left( \frac{n}{n+1} \right)^{\epsilon_n} = \frac{1}{2} \prod_{n=1}^{\infty} \left( \frac{2n}{2n+1} \right)^{\epsilon_{2n}} \cdot \prod_{n=0}^{\infty} \left( \frac{2n+1}{2n+2} \right)^{\epsilon_{2n+1}}$$

$$\epsilon_{2n} = (-1)^{t_{2n}} = (-1)^{t_n} = \epsilon_n \quad \epsilon_{2n+1} = (-1)^{t_{2n+1}} = (-1)^{1-t_n} = -\epsilon_n$$

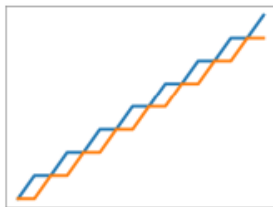
$$PQ = \frac{1}{2} \prod_{n=1}^{\infty} \left( \frac{2n}{2n+1} \right)^{\epsilon_n} \cdot \prod_{n=0}^{\infty} \left( \frac{2n+1}{2n+2} \right)^{-\epsilon_n} = \frac{1}{2} \prod_{n=1}^{\infty} \left( \frac{2n}{2n+1} \right)^{\epsilon_n} \cdot \left( \prod_{n=0}^{\infty} \left( \frac{2n+1}{2n+2} \right)^{\epsilon_n} \right)^{-1} = \frac{1}{2} \frac{Q}{P}$$

$$PQ = \frac{1}{2} \frac{Q}{P} \rightarrow P^2 = \frac{1}{2} \rightarrow P = \sqrt{\frac{1}{2}} = \frac{\sqrt{2}}{2}$$

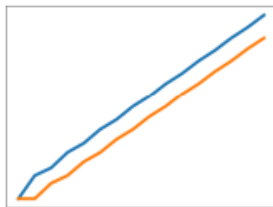
Q=?

ONE **COVRILUCA** FOR THOSE  
WHO CAN GIVE ME THE ANSWER

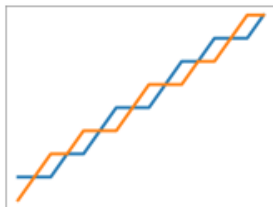
# Fair Share Sequence



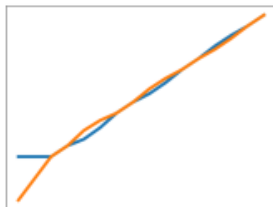
(a) *ABABABABAB...*



(b) Running average



(c) Thue-Morse



(d) Thue-Morse average

# Fair Share Sequence

Darius and Mihai are dividing things of non-increasing value amongst themselves. What is the fairest order for them to pick?

# Fair Share Sequence

Darius and Mihai are dividing things of non-increasing value amongst themselves. What is the fairest order for them to pick?

Suppose Darius picks first. Then Mihai should pick second, then Darius picks third, Mihai picks fourth, etc:

DMDMDMDMDM ...

# Fair Share Sequence

Darius and Mihai are dividing things of non-increasing value amongst themselves. What is the fairest order for them to pick?

Suppose Darius picks first. Then Mihai should pick second, then Darius picks third, Mihai picks fourth, etc:

DMDMDMDMDM ...

Darius gets an advantage: For every pair of items, Darius will get to pick the better one!

# Fair Share Sequence

Maybe after DM, what if they swapped order after?

DM MD

# Fair Share Sequence

Maybe after DM, what if they swapped order after?

DM MD

Now it's more fair if there are 4 items, but if we repeat this:

DMMD DMMD DMMD DMMD ...



# Fair Share Sequence

Maybe after DM, what if they swapped order after?

DM MD

Now it's more fair if there are 4 items, but if we repeat this:

DMMD DMMD DMMD DMMD ...

Darius gets an advantage again: Darius will get to pick the best item out of every 4 items!

# Fair Share Sequence

Let's flip the order again:

DMMD MDDM

# Fair Share Sequence

Let's flip the order again:

DMMD MDDM

Again, if we repeat this, Darius will get to pick the best item out of every 8 items!

# Fair Share Sequence

Let's flip the order again:

DMMD MDDM

Again, if we repeat this, Darius will get to pick the best item out of every 8 items!

If we keep flipping the order,

DMMD MDDM MDDM DMMD MDDM DMMD DMMD MDDM ...

If we replace  $D \rightarrow 0$  and  $M \rightarrow 1$ , this is the Thue-Morse sequence!

# Magic Squares

By a magic square of order  $T$ , we mean a  $T$  by  $T$  matrix whose entries are taken from the numbers  $1, 2, 3, \dots, T^2$ , and such that the sum of the entries in any row, column, or diagonal is the same number.

One also requires that each of the numbers  $1, 2, 3, \dots, T^2$  be used exactly once as an entry.

## Example

16	3	2	13
5	10	11	8
9	6	7	12
4	15	14	1

# Construction of magic squares of order $T = 2^M$ with special properties using Thue-Morse Sequence

We illustrate this by constructing a  $4 \times 4$  magic square, i.e. putting  $M = 2$ :

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$t =$	0	1	1	0	1	0	0	1	1	0	0	1	0	1	1	0
	$t_0$	$t_1$	$t_2$	$t_3$	$t_4$	$t_5$	$t_6$	$t_7$	$t_8$	$t_9$	$t_{10}$	$t_{11}$	$t_{12}$	$t_{13}$	$t_{14}$	$t_{15}$

16	2	3	13
5	11	10	8
9	7	6	12
4	14	15	1

The  $n$ -th box of the magic square is  $n$  if  $t_{n-1} = 1$ . The unused numbers are inserted in decreasing order into the empty boxes

# Magic Square

## Remark

The sum of the entries in any row, column or diagonal of a magic square of order  $T$  is

$$\frac{1}{2}T(T^2 + 1)$$

## Facts

- 1 It is known that one can construct magic squares of any order except 2.
- 2 Any magic square can be rotated and reflected to produce 8 trivially distinct squares, and the eight such squares are said to make up a single equivalence class.
- 3 There is exactly one  $3 \times 3$  magic square, exactly 880  $4 \times 4$  magic squares and exactly 275,305,224  $5 \times 5$  magic squares.

# Pattern Avoidance

An **overlap** is a string of the form  $axaxa$  where  $a$  is a single letter and  $x$  is a string.

Pick a binary word  $w$ , such as  $w = 1101$ .

Consider blocks of the form  $0w0w0$  and  $1w1w1$ .

Neither block appears ( no matter which  $w$  is chosen ).

The Thue-Morse Sequence is *overlap free*.

011010011001011010010110011010011001011001101001011010...

01101011010

1110111011



# Pattern Avoidance

Count the numbers of 1s between successive 0s.

011010011001011010010110011010011001011001101001011010...

2

# Pattern Avoidance

Count the numbers of 1s between successive 0s.

011010011001011010010110011010011001011001101001011010...

21

# Pattern Avoidance

Count the numbers of 1s between successive 0s.

011010011001011010010110011010011001011001101001011010...

210

# Pattern Avoidance

Count the numbers of 1s between successive 0s.

011010011001011010010110011010011001011001101001011010...

2102

# Pattern Avoidance

We keep doing this counting the number of 1s between the next two successive 0s until we get an infinite integer sequence made of 2s, 1s and 0s like this one :

210201210120210201202102012021012102012102012102012...

A **square** is a string of the form  $xx$ .

A word is **squarefree** if it contains no subword (block of consecutive symbols) that is a square. Note that **squarefree** is not squarefree, but **square** is.

Consider any ternary (base 3) word  $w$  such as  $w = 01212$ . If we look through the entire sequence to see if we can find this block 0121201212 it turns out that no matter which  $w$  we choose the repeated block pattern  $ww$  will never appear. This new sequence is called *square free*.

210201210120210201202102012021012102012102012102012...

0121201212

Thue-Morse Sequence has even been used to composing music!

Tom Johnson, a Paris-based composer, has used the Thue-Morse sequence and other sequences formed by iterated morphisms, in his work.



Figure: Composition by Tom Johnson

# Twenty Questions with a liar

Consider a game in which one player chooses an integer  $x \in [1, 1000]$ , and the other player has to guess it by asking yes/no questions. The second player is allowed to give at most one incorrect answer.

What is the minimum number of yes/no questions required to guarantee the identification of the chosen number?

30...27...21...who says lower? Place your bets!

# Twenty Questions with a liar - Solution

Let us denote the bits of the first player's number by  $a_i$ , where  $i \in [1..10]$ . We start by asking the values of all those bits. That is, we ask the following questions: "Is it true that the  $i$ th bit of your number is zero?" Let us denote the answers to these questions as  $b_i$ , where  $i \in [1..10]$ .

<u>        </u>	<u>        </u>	<u>        </u>	<u>        </u>	<u>        </u>	<u>        </u>	<u>        </u>	<u>        </u>	<u>        </u>	<u>        </u>
$a_{10}$	$a_9$	$a_8$	$a_7$	$a_6$	$a_5$	$a_4$	$a_3$	$a_2$	$a_1$
<u>        </u>	<u>        </u>	<u>        </u>	<u>        </u>	<u>        </u>	<u>        </u>	<u>        </u>	<u>        </u>	<u>        </u>	<u>        </u>
$b_{10}$	$b_9$	$b_8$	$b_7$	$b_6$	$b_5$	$b_4$	$b_3$	$b_2$	$b_1$



# Twenty Questions with a liar - Solution

Now, we ask 4 additional questions:

- Is it true that  $a_1 \oplus a_2 \oplus a_4 \oplus a_5 \oplus a_7 \oplus a_9 = 0$ ? ( $\oplus$  denotes summation modulo 2.)
- Is it true that  $a_1 \oplus a_3 \oplus a_4 \oplus a_6 \oplus a_7 \oplus a_{10} = 0$ ?
- Is it true that  $a_2 \oplus a_3 \oplus a_4 \oplus a_8 \oplus a_9 \oplus a_{10} = 0$ ?
- Is it true that  $a_5 \oplus a_6 \oplus a_7 \oplus a_8 \oplus a_9 \oplus a_{10} = 0$ ?

## Twenty Questions with a liar - Solution

Let  $q_1, q_2, q_3$ , and  $q_4$  be the answers to these additional questions. Now, the second player calculates  $t_i$  for  $i \in [1..4]$  — answers to those questions based on the bits  $b_j$  which he previously received from the first player. There are 16 ways in which the bits  $q_i$  can differ from  $t_i$ . Let  $d_i = q_i \oplus t_i$  (thus  $d_i = 1$  if and only if  $q_i \neq t_i$ ).

# Twenty Questions with a liar - Solution

Let us now construct a table of all possible errors and corresponding values of  $(d_1, d_2, d_3, d_4)$ :

Error in	$(d_1, d_2, d_3, d_4)$
No error	$(0,0,0,0)$
$b_1$	$(1,1,0,0)$
$b_2$	$(1,0,1,0)$
$b_3$	$(0,1,1,0)$
$b_4$	$(1,1,1,0)$
$b_5$	$(1,0,0,1)$
$b_6$	$(0,1,0,1)$
$b_7$	$(1,1,0,1)$
$b_8$	$(0,0,1,1)$
$b_9$	$(1,0,1,1)$
$b_{10}$	$(0,1,1,1)$

Error in	$(d_1, d_2, d_3, d_4)$
$q_1$	$(1,0,0,0)$
$q_2$	$(0,1,0,0)$
$q_3$	$(0,0,1,0)$
$q_4$	$(0,0,0,1)$

All values of  $(d_1, d_2, d_3, d_4)$  are distinct, so the error can be identified and corrected.