

A LINEAR-SIZE BLOCK-PARTITION FIBONACCI ENCODING FOR GÖDEL NUMBERING

ZOLTÁN SÓSTAI

ABSTRACT. We construct an encoding of finite strings over a fixed finite alphabet as natural numbers, based on a block partition of the Fibonacci sequence. Each position in the string selects one Fibonacci number from a dedicated block, with unused indices between blocks guaranteeing non-adjacency. The encoded number is the sum of the selected Fibonacci numbers, and Zeckendorf’s theorem [6, 4] guarantees that this sum uniquely determines the selection. The encoding is injective, the string length is recoverable from the code, and the worst-case digit count of the encoded number grows as $\Theta(m)$ for strings of length m —matching the information-theoretic lower bound up to a constant factor. We also prove that the natural right-nested use of Rosko’s [5] binary carryless pairing for sequence encoding has worst-case $\Theta(2^m)$ digit growth—an exponential blowup that the block-partition construction avoids entirely.

1. Introduction

Gödel’s incompleteness theorems [2] require encoding syntactic objects—symbols, formulas, proofs—as natural numbers. The original encoding maps a sequence $\langle a_1, \dots, a_m \rangle$ to the product

$$2^{a_1+1} \cdot 3^{a_2+1} \cdot 5^{a_3+1} \dots p_m^{a_m+1},$$

where p_i is the i -th prime and the $+1$ ensures positive exponents. The fundamental theorem of arithmetic guarantees injectivity, and decoding reduces to factorization. This encoding produces very large numbers: even short formulas yield codes of dozens or hundreds of digits.

Rosko [5] proposed an alternative based on Zeckendorf representations—unique decompositions of natural numbers as sums of non-consecutive Fibonacci numbers [6, 4]. His “carryless pairing” function is a pairing $\mathbb{N}^2 \rightarrow \mathbb{N}$ that places the Zeckendorf supports of two inputs into disjoint even/odd index bands. Encoding sequences of length m via right-nesting of this binary pairing requires $m - 1$ applications. We prove in §5.3 that this right-nested use has worst-case $\Theta(2^m)$ digit growth.

The present paper introduces a simpler construction: a **block-partition encoding** that partitions the Fibonacci sequence into contiguous blocks, one per string position, with single-index gaps between blocks to ensure non-adjacency. Each symbol selects one Fibonacci number from its position’s block, and the encoded number is the sum. No nesting is involved. The digit count of the encoded number grows linearly in the string length.

The paper isolates the **numeric coding layer** of a Gödel numbering: the injective mapping from strings to natural numbers. The further machinery required for a full syntactic arithmetization—substitution, concatenation, diagonalization—is not addressed here.

Remark 1.1. The term “Fibonacci coding” already denotes a variable-length binary code used in data compression [1], where each positive integer is represented by a binary string derived from its Zeckendorf representation. Our construction is distinct: we encode *strings over a formal alphabet* as single natural numbers. The shared foundation is Zeckendorf’s theorem, but the purpose and mechanism differ.

Date: March 2026.

2020 Mathematics Subject Classification. 03F30, 11B39, 03D20.

Key words and phrases. Zeckendorf representation, Fibonacci encoding, Gödel numbering, block partition.

2. Preliminaries

2.1. Fibonacci sequence. Let $\{F_n\}_{n \geq 0}$ denote the Fibonacci sequence defined by $F_0 = 0$, $F_1 = 1$, and $F_{n+2} = F_{n+1} + F_n$ for all $n \geq 0$. The sequence begins

$$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, 987, \dots$$

It is well known that $F_n = (\varphi^n - \psi^n)/\sqrt{5}$, where $\varphi = (1 + \sqrt{5})/2 \approx 1.618$ is the golden ratio and $\psi = (1 - \sqrt{5})/2 \approx -0.618$ (see, e.g., [3], §6.6).

2.2. Zeckendorf's theorem.

Theorem 2.1 (Zeckendorf [6], Lekkerkerker [4]). *Every positive integer n has a unique representation as a sum of non-consecutive Fibonacci numbers:*

$$n = F_{e_1} + F_{e_2} + \dots + F_{e_s}$$

where $e_1 > e_2 > \dots > e_s \geq 2$ and $e_i - e_{i+1} \geq 2$ for all i .

The set $Z(n) = \{e_1, e_2, \dots, e_s\}$ is called the **Zeckendorf support** of n . We define $Z(0) = \emptyset$. The support of a positive integer is computed by the **greedy algorithm**: at each step, select the largest Fibonacci number not exceeding the remainder, subtract it, and repeat.

3. The block-partition encoding

3.1. The alphabet. Let $\Sigma = \{\sigma_0, \sigma_1, \dots, \sigma_{k-1}\}$ be a finite alphabet of $k \geq 2$ symbols. Each symbol σ_j is assigned the **symbol code** $j \in \{0, 1, \dots, k-1\}$.

For illustration, we use a 10-symbol alphabet drawn from Peano arithmetic:

Code	0	1	2	3	4	5	6	7	8	9
Symbol	0	<i>S</i>	+	·	=	¬	→	∀	()

A complete first-order syntax would require a finite tokenization scheme for variables (e.g., de Bruijn indices), proof delimiters, and possibly additional connectives. The construction handles any finite alphabet identically by adjusting k .

3.2. Block partition.

Definition 3.1 (Block partition). For a k -symbol alphabet, define the **block size** $\beta = k + 1$. For each position $i \in \{0, 1, 2, \dots\}$ in a string, define the **position block** B_i as the set of Fibonacci indices:

$$B_i = \{2 + i\beta, 2 + i\beta + 1, \dots, 2 + i\beta + (k - 1)\}.$$

Each block B_i contains exactly k consecutive Fibonacci indices—one for each symbol. The index $2 + i\beta + k$ (the first index of the gap) is **unused**, ensuring that the last index of B_i and the first index of B_{i+1} are separated by at least 2.

For the 10-symbol alphabet ($k = 10$, $\beta = 11$), the encoding is displayed in Table 1. Each row corresponds to one symbol and shows the Fibonacci sequence from F_2 onward. **Bold** numbers are the Fibonacci numbers assigned to that symbol at each successive position; *italicized* numbers are the unused gap entries. The three blocks shown correspond to positions 1, 2, and 3.

	Block B_0 (Pos.1)										Block B_1 (Pos.2)										Block B_2 (Pos.3)												
:																																	
0:	1	2	3	5	8	13	21	34	55	89	<i>144</i>	233	377	610	987	1597	2584	4181	6765	10946	17711	<i>28657</i>	46368	75025	121393	196418	317811	514229	832040	1346269	2178309	3524578	<i>5702887</i>
S:	1	2	3	5	8	13	21	34	55	89	<i>144</i>	233	377	610	987	1597	2584	4181	6765	10946	17711	<i>28657</i>	46368	75025	121393	196418	317811	514229	832040	1346269	2178309	3524578	<i>5702887</i>
+	1	2	3	5	8	13	21	34	55	89	<i>144</i>	233	377	610	987	1597	2584	4181	6765	10946	17711	<i>28657</i>	46368	75025	121393	196418	317811	514229	832040	1346269	2178309	3524578	<i>5702887</i>
·	1	2	3	5	8	13	21	34	55	89	<i>144</i>	233	377	610	987	1597	2584	4181	6765	10946	17711	<i>28657</i>	46368	75025	121393	196418	317811	514229	832040	1346269	2178309	3524578	<i>5702887</i>
=	1	2	3	5	8	13	21	34	55	89	<i>144</i>	233	377	610	987	1597	2584	4181	6765	10946	17711	<i>28657</i>	46368	75025	121393	196418	317811	514229	832040	1346269	2178309	3524578	<i>5702887</i>
→	1	2	3	5	8	13	21	34	55	89	<i>144</i>	233	377	610	987	1597	2584	4181	6765	10946	17711	<i>28657</i>	46368	75025	121393	196418	317811	514229	832040	1346269	2178309	3524578	<i>5702887</i>
∨	1	2	3	5	8	13	21	34	55	89	<i>144</i>	233	377	610	987	1597	2584	4181	6765	10946	17711	<i>28657</i>	46368	75025	121393	196418	317811	514229	832040	1346269	2178309	3524578	<i>5702887</i>
(1	2	3	5	8	13	21	34	55	89	<i>144</i>	233	377	610	987	1597	2584	4181	6765	10946	17711	<i>28657</i>	46368	75025	121393	196418	317811	514229	832040	1346269	2178309	3524578	<i>5702887</i>
)	1	2	3	5	8	13	21	34	55	89	<i>144</i>	233	377	610	987	1597	2584	4181	6765	10946	17711	<i>28657</i>	46368	75025	121393	196418	317811	514229	832040	1346269	2178309	3524578	<i>5702887</i>

TABLE 1. Block-partition mapping for the 10-symbol alphabet. Each row shows the Fibonacci sequence for one symbol, with **bold** entries marking the assigned Fibonacci number at each position (blocks of 10, separated by italicized gap entries). To encode a string: for each position, find the row of the symbol at that position, read off its bold entry for that block, and sum all selected entries. Zeckendorf's theorem guarantees that the sum uniquely determines the selection.

3.3. Encoding.

Definition 3.2 (Block-partition encoding). Let $s = s_1 s_2 \cdots s_m$ be a nonempty string of length m over Σ , with symbol codes $c_1, c_2, \dots, c_m \in \{0, 1, \dots, k-1\}$. The **block-partition Fibonacci code** of s is

$$\lceil s \rceil = \sum_{i=0}^{m-1} F_{2+i\beta+c_{i+1}}.$$

That is, for position i (0-indexed), select the Fibonacci number at index $2+i\beta+c_{i+1}$ from block B_i . For the empty string, define $\lceil \varepsilon \rceil = 0$.

Example 3.3. The string $0 = 0$ has symbol codes $(0, 4, 0)$. Under the encoding:

Position 0, symbol 0 (code 0): $F_2 = 1$

Position 1, symbol = (code 4): $F_{17} = 1597$

Position 2, symbol 0 (code 0): $F_{24} = 46368$

Hence $\lceil 0 = 0 \rceil = 1 + 1597 + 46368 = \mathbf{47966}$.

Example 3.4. The string $S(0) = S(0)$ has symbol codes $(1, 8, 0, 9, 4, 1, 8, 0, 9)$. Under the encoding:

Position 0, symbol S (code 1): $F_3 = 2$

Position 1, symbol ((code 8): $F_{21} = 10946$

Position 2, symbol 0 (code 0): $F_{24} = 46368$

Position 3, symbol) (code 9): $F_{44} = 701408733$

Position 4, symbol = (code 4): $F_{50} = 12586269025$

Position 5, symbol S (code 1): $F_{58} = 591286729879$

Position 6, symbol ((code 8): $F_{76} = 3416454622906707$

Position 7, symbol 0 (code 0): $F_{79} = 14472334024676221$

Position 8, symbol) (code 9): $F_{99} = 218922995834555169026$

Hence $\lceil S(0) = S(0) \rceil = 21894088522777216907$, a 21-digit number. For comparison, the prime-power Gödel encoding of the same string produces a number of approximately 49 digits.

4. Correctness

4.1. Non-adjacency.

Lemma 4.1 (Non-adjacency). *For any nonempty string $s = s_1 s_2 \cdots s_m$ over a k -symbol alphabet, the Fibonacci indices selected by the block-partition encoding are mutually non-consecutive.*

Proof. Each position i selects exactly one index from block B_i . Within a single block, only one index is selected, so no adjacency arises within a block.

Between consecutive blocks: the maximum index in B_i is $2 + i\beta + (k - 1)$. The minimum index in B_{i+1} is $2 + (i + 1)\beta = 2 + i\beta + k + 1$. The difference is

$$(2 + i\beta + k + 1) - (2 + i\beta + k - 1) = 2,$$

so the two indices are non-consecutive in the Fibonacci sequence. For non-adjacent blocks B_i and B_j with $j > i + 1$, the gap is strictly larger. \square

4.2. Injectivity.

Theorem 4.2 (Injectivity). *The block-partition encoding is injective on Σ^* : if $s \neq t$, then $\lceil s \rceil \neq \lceil t \rceil$.*

Proof. If one of s, t is the empty string and the other is not, then one code is 0 and the other is a positive integer (since every nonempty string selects at least one $F_e \geq F_2 = 1$).

If s and t are both nonempty with $|s| = m < m' = |t|$, then the Zeckendorf support of $\lceil t \rceil$ contains an index from block B_m or higher, while every index in $Z(\lceil s \rceil)$ lies in blocks B_0 through B_{m-1} . These are distinct subsets of non-consecutive indices (by Lemma 4.1), so by Theorem 2.1 the sums differ.

If $|s| = |t|$ but s and t differ in at least one position j , then position j selects a different Fibonacci index in s than in t . The remaining positions select identical indices. So the two Zeckendorf supports differ, and by Theorem 2.1 the sums differ. \square

4.3. Valid codes and length recovery.

Definition 4.3 (Valid code). A natural number n is a **valid code** if $n = 0$ (encoding the empty string) or if there exist $m \geq 1$ and unique $c_1, c_2, \dots, c_m \in \{0, 1, \dots, k - 1\}$ such that

$$Z(n) = \{ 2 + i\beta + c_{i+1} : 0 \leq i < m \}.$$

That is, the Zeckendorf support of n consists of exactly m indices, one from each block B_0 through B_{m-1} , with no indices in gap positions. The image of the encoding $\lceil \cdot \rceil$ is exactly the set of valid codes.

Lemma 4.4 (Length recovery). *For any valid code n , the length of the encoded string is recoverable: if $n = 0$, then $m = 0$; if $n > 0$, then $m = |Z(n)|$.*

Proof. For $n > 0$, Definition 4.3 requires $Z(n)$ to consist of exactly m indices, one per block. Therefore $|Z(n)| = m$. Equivalently, $m = \lfloor (e_{\max} - 2)/\beta \rfloor + 1$ where $e_{\max} = \max Z(n)$, since the largest index lies in B_{m-1} . \square

4.4. Decoding.

Theorem 4.5 (Decoding). *For any valid code n , the original string s with $\lceil s \rceil = n$ is uniquely recoverable.*

Proof. If $n = 0$, return the empty string. Otherwise, apply the greedy Zeckendorf algorithm to n to obtain the support $Z(n)$. By Definition 4.3 and Lemma 4.4, this support contains exactly m indices, one per block. For each recovered index e_j , compute

$$\text{position: } i = \lfloor (e_j - 2)/\beta \rfloor, \quad \text{symbol code: } c = (e_j - 2) \bmod \beta.$$

This recovers the symbol at each position. The division and modulus by the fixed constant $\beta = k + 1$ are computable by repeated subtraction. \square

5. Complexity analysis

5.1. Code length.

Theorem 5.1 (Linear code length). *For $k \geq 2$ and $m \geq 1$, the worst-case digit count $L(m) = \max\{\text{digits}(\lceil s \rceil) : s \in \Sigma^m\}$ satisfies $L(m) = \Theta(m)$.*

Proof. The largest Fibonacci index used in encoding a string of length m is at most

$$e_{\max} = 2 + (m - 1)\beta + (k - 1) = m\beta.$$

Every length- m code satisfies

$$F_{2+(m-1)\beta} \leq \lceil s \rceil < F_{m\beta+1}.$$

The lower bound holds because the code includes at least $F_{2+(m-1)\beta}$ from block B_{m-1} . The upper bound holds because the sum of any set of non-consecutive Fibonacci numbers with indices up to e is strictly less than F_{e+1} (a standard property of Zeckendorf representations).

Taking logarithms, both bounds are $\Theta(m\beta \cdot \log_{10} \varphi) = \Theta(m)$. For the 10-symbol alphabet ($\beta = 11$), this gives approximately 2.3 digits per symbol. \square

5.2. Optimality.

Theorem 5.2 (Asymptotic optimality). *For $k \geq 2$, the block-partition encoding has asymptotically optimal code length: $L(m) = \Theta(m)$, matching the information-theoretic lower bound.*

Proof. There are k^m distinct strings of length m . Any injective map into \mathbb{N} must use numbers as large as $k^m - 1$, requiring at least $m \cdot \log_{10}(k)$ digits. The block-partition encoding achieves $L(m) = \Theta(m \cdot (k + 1) \cdot \log_{10} \varphi)$. The ratio of the constant factors is

$$\frac{(k + 1) \log_{10} \varphi}{\log_{10} k}.$$

For $k = 10$, this is $11 \cdot 0.209/1 \approx 2.3$. Both bounds are linear in m ; the block-partition encoding incurs a constant-factor overhead of approximately 2.3 relative to an information-theoretically optimal encoding. \square

5.3. Exponential blowup of right-nested carryless pairing. Rosko [5] defines a carryless pairing function $\pi : \mathbb{N}^2 \rightarrow \mathbb{N}$. His paper presents it as a binary pairing and sketches a k -ary generalization, but does not develop a completed sequence encoding. The natural way to encode a length- m sequence (a_1, a_2, \dots, a_m) using a binary pairing is right-nesting. Formally, define $t_1 = a_m$ and $t_{j+1} = \pi(a_{m-j}, t_j)$ for $j = 1, \dots, m - 1$. The final code is t_{m-1} , requiring $m - 1$ pairings. Let $M_j = \max Z(t_j)$ denote the largest Fibonacci index in the support after j pairings.

To avoid the issue that $Z(0) = \emptyset$ (making Rosko's pairing degenerate for zero inputs—in particular $\pi(0, 0) = 0$), we adopt the standard shift-by-one convention: symbol codes are drawn from $\{1, 2, \dots, k\}$ rather than $\{0, \dots, k - 1\}$.

Proposition 5.3 (Exponential growth). *Let codes be drawn from $\{1, \dots, k\}$. Right-nested encoding of a length- m sequence via Rosko's carryless pairing produces codes whose worst-case digit count is $\Theta(2^m)$.*

Proof. In Rosko's pairing $\pi(x, y)$, the output's Zeckendorf support occupies indices up to $B(x) + 2R(y) - 1$, where $B(x) = 2r(x)$ is the delimiter, $r(x) = \min\{e : F_e > x\}$ is the rank, and $R(y) = \max Z(y)$. Since all symbol codes are positive, $Z(a_i)$ is nonempty and $R(a_i)$ is well-defined for every symbol.

Over the fixed symbol set $\{1, \dots, k\}$, define

$$b_{\min} = 2 \cdot \min\{r(a) : a \in \{1, \dots, k\}\}, \quad b_{\max} = 2 \cdot \max\{r(a) : a \in \{1, \dots, k\}\}.$$

At each nesting step, the odd band determines the growth:

$$2M_j + b_{\min} - 1 \leq M_{j+1} \leq 2M_j + b_{\max} - 1.$$

Both bounds are linear recurrences of the form $M_{j+1} = 2M_j + C$, solving to $A \cdot 2^j - C$ for a constant A determined by initial conditions. Therefore $M_j = \Theta(2^j)$. Since the digit count is $\Theta(M_j \cdot \log_{10} \varphi)$, the encoded number has $\Theta(2^j)$ digits. A length- m sequence requires $m - 1$ pairings, giving worst-case $\Theta(2^m)$ digit growth. \square

For concreteness: with symbol code 5 (where $r(5) = 6$, $B(5) = 12$, $R(5) = 5$), the recurrence gives $M_1 = 21$, $M_2 = 53$, $M_3 = 117$, $M_4 = 245$. A length-5 sequence requires 4 pairings, so the largest Fibonacci index is $M_4 = 245$. By comparison, the block-partition encoding of a length-5 string uses a maximum Fibonacci index of $5 \cdot 11 = 55$.

6. Comparison of encodings

Encoding	Code length growth
Prime-power (Gödel, 1931)	Super-linear
Base- k positional	Linear (optimal constant)
Right-nested carryless pairing (Rosko, 2025)	Exponential: $\Theta(2^m)$ (Prop. 5.3)
Block partition (this paper)	Linear: $\Theta(m)$ (Thm. 5.1)

TABLE 2. Worst-case code length for a string of length m . To our knowledge, the block-partition construction is the first Zeckendorf-based encoding to achieve linear code growth.

7. Conclusion

The block-partition Fibonacci encoding provides a simple, compact mapping from finite strings to natural numbers, with worst-case code length growing linearly in the string length. The construction rests on an elementary observation: partitioning the Fibonacci sequence into blocks with single-index gaps ensures that any selection of one Fibonacci number per block yields a valid Zeckendorf representation, and Zeckendorf’s theorem then guarantees unique decodability.

For sequence encoding over the same Zeckendorf foundation, the block-partition construction achieves $\Theta(m)$ code length where the natural right-nested use of Rosko’s binary carryless pairing produces $\Theta(2^m)$ code length (Proposition 5.3), an exponential improvement.

Whether the additive character of this encoding can be exploited for theory-internal purposes—such as establishing Δ_0 -definability of syntactic predicates in bounded arithmetic or enabling incompleteness proofs in systems lacking multiplication—is a question we leave to future work.

References

- [1] A. S. Fraenkel, Systems of numeration, *Amer. Math. Monthly* **92** (1985), no. 2, 105–114.
- [2] K. Gödel, Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I, *Monatsh. Math. Phys.* **38** (1931), no. 1, 173–198.
- [3] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics: A Foundation for Computer Science*, 2nd ed., Addison-Wesley, 1994.
- [4] C. G. Lekkerkerker, Voorstelling van natuurlijke getallen door een som van getallen van Fibonacci, *Simon Stevin* **29** (1952), 190–195.
- [5] M. Rosko, Carryless pairing: Additive pairing in the Fibonacci basis, *arXiv:2509.10382*, 2025.
- [6] E. Zeckendorf, Représentation des nombres naturels par une somme de nombres de Fibonacci ou de nombres de Lucas, *Bull. Soc. Roy. Sci. Liège* **41** (1972), no. 4–6, 179–182.