

# A General Approach to Proving Properties of Fibonacci Representations via Automata Theory

Jeffrey Shallit\* and Sonja Linghui Shan

School of Computer Science, University of Waterloo, Waterloo, ON N2L 3G1, Canada

shallit@uwaterloo.ca, slshan@uwaterloo.ca

We provide a method, based on automata theory, to mechanically prove the correctness of many numeration systems based on Fibonacci numbers. With it, long case-based and induction-based proofs of correctness can be replaced by simply constructing a regular expression (or finite automaton) specifying the rules for valid representations, followed by a short computation. Examples of the systems that can be handled using our technique include Brown’s lazy representation (1965), the far-difference representation developed by Alpert (2009), and three representations proposed by Hajnal (2023). We also provide three additional systems and prove their validity.

## 1 Introduction

Given an increasing sequence  $(s_n)_{n \geq 0}$  of positive integers, a numeration system is a way of expressing natural numbers as a linear combination of the  $s_n$ . Many different numeration systems, such as representation in base  $k$ , or the more exotic systems based on the Fibonacci numbers, have been proposed. For example, recall that the Fibonacci numbers, sequence [A000045](#) in the *On-Line Encyclopedia of Integer Sequences* (OEIS), are defined by the recurrence  $F_n = F_{n-1} + F_{n-2}$  for  $n \geq 2$  and the initial values  $F_0 = 0$ ,  $F_1 = 1$ . Consider writing a non-negative integer  $n$  as a sum of distinct Fibonacci numbers  $F_i$  for  $i \geq 2$ . Some numbers, such as 12, have only one such representation ( $12 = 8 + 3 + 1 = F_6 + F_4 + F_2$ ), while others have many:  $8 = F_6 = F_5 + F_4 = F_5 + F_3 + F_2$ .

There are two very desirable characteristics of a numeration system. First, *completeness*: every natural number should have a representation. Second, *unambiguity*: no natural number should have two or more different representations. These two goals are typically achieved by restricting the types of representations that are considered valid within the system. If a system achieves both goals, we say it is *perfect*. For Fibonacci representations, various perfect systems have been proposed.

Among all possible perfect systems based on Fibonacci numbers, one is particularly useful: the *Zeckendorf* or *greedy* representation. This representation can be computed as follows: first, choose the largest index  $i$  such that  $F_i \leq n$ . Then the representation for  $n$  is  $F_i$  plus the (recursively-computed) representation for  $n - F_i$ . The representation for 0 is the empty sum of 0 Fibonacci numbers. A simple induction now shows that the greedy algorithm produces a representation for every natural number, which is evidently unique.

This representation was originally noted by Zeckendorf, but was first published by Lekkerkerker [12] and only later by Zeckendorf himself [20]. It was also anticipated, in much more general form, by Ostrowski [15].

An alternative (but equivalent) definition of Zeckendorf representation is to impose a condition that valid representations must obey. For example, we could require that a representation be valid if and only if no two consecutive Fibonacci numbers appear in the sum.

---

\*Research funded by a grant from NSERC, 2018-04118.

It is convenient to express arbitrary sums of distinct Fibonacci numbers as strings of digits over a finite alphabet (in analogy with base- $k$  representation). Let  $x = a_1 \cdots a_t$  be a string (or word) made up of integer digits. We define its value as a Fibonacci representation as follows:

$$[x]_F := \sum_{1 \leq i \leq t} a_i F_{t+2-i}. \tag{1}$$

Note that these strings are in “most-significant-digit” first format. For example,  $[2101]_F = 2F_5 + F_4 + F_2 = 14$ .

It is also useful to define a (partial) inverse to  $[x]_F$ . By  $(n)_F$  we mean the binary string  $x$  such that  $x$  is the Zeckendorf representation of  $n$ ; alternatively, such that  $[x]_F = n$  and  $x$  contains no occurrence of the block 11. In what follows, we adopt this string-based point of view almost exclusively. We can think of the condition “no occurrence of the block 11” as a *rule*, specifying which representations are valid, adopted precisely to guarantee both completeness and unambiguity.

In formal language theory, a language  $L$  is a (finite or infinite) collection of strings. A rule is then encoded by the language or set of strings that obey the rule. Completeness then becomes the assertion that for all  $n$  there exists a string  $x \in L$  such that  $[x]_F = n$ , while unambiguity becomes the assertion that there do not exist distinct strings  $x, y \in L$  such that  $[x]_F = [y]_F$ .<sup>1</sup>

Let us look at another example involving the Fibonacci numbers, one that is much less well known: the so-called *lazy* representation [3]. In this system, representation as a sum of Fibonacci numbers corresponds (via Eq. (1)) to a binary string having no occurrence of the block 00 (where leading zeros are not even considered). Once again, this rule provides a numeration system that is both complete and unambiguous [3]. Table 1 gives both greedy (Zeckendorf) and lazy representations for the first few natural numbers.

$n$	0	1	2	3	4	5	6	7	8	9	10	11
greedy	$\varepsilon$	1	10	100	101	1000	1001	1010	10000	10001	10010	10100
lazy	$\varepsilon$	1	10	11	101	110	111	1010	1011	1101	1110	1111

Table 1: Greedy and lazy Fibonacci representations.

The greedy and lazy representations are certainly not the only possible perfect numeration systems based on the Fibonacci numbers. In fact, there are *uncountably many* such systems! These result from making a choice, for all  $n$  having at least two different representations as sums of distinct Fibonacci numbers, about which particular representation is chosen to be valid. (By a result of Robbins [17], “most” numbers have more than one representation as a sum of distinct Fibonacci numbers.)

If we demand that the set of valid representations forms a regular language—that is, accepted by a finite automaton; see Section 2—there are still infinitely many different systems (although only countably many). For example, consider choosing the  $t$ ’th largest possible representation for  $n$  in lexicographic order (if there are at least  $t$ ), and otherwise the lexicographically first. It will follow from results below that, for each  $t \geq 0$ , this choice gives a regular language  $L_t$  of valid representations.

Some natural questions then arise: given a language  $L$  encoding the “rule” a representation must obey (such as no occurrence of the block 11, or no occurrence of the block 00), how can we determine if the corresponding set of Fibonacci representations is complete and unambiguous? And if it is complete,

<sup>1</sup>We adopt the convention that two strings are considered to be the same if they differ only in the number of leading zeros. Thus, for example,  $[100]_F = [0100]_F = 3$  are the same representation.

how can we efficiently find a representation for a given number  $n$ ? Up to now, each new system proposed required a new proof, often a rather tedious case-based proof by induction. In this paper we provide a *general framework* for answering these questions “automatically”, via an algorithm, in the case where the language of valid representations is regular.

These ideas are capable of generalization. For example, we can also consider representations for all integers  $\mathbb{Z}$ , instead of just the natural numbers  $\mathbb{N}$ . This can be achieved in two distinct ways:

- By allowing a larger digit set, say,  $\{-1, 0, 1\}$ ;
- By using the so-called *negaFibonacci* system, based on the Fibonacci numbers of negative index  $F_{-n}$  for  $n \geq 1$ .

Once again, we would like a choice of valid representations that is complete and unambiguous.

In this paper we show how to decide these properties, provided that the set of valid representations forms a regular language (which is indeed the case for all the proposed systems in the literature).

Here is an outline of the paper. In Section 2, we explain the basics of automata theory needed to understand the rest of the paper. In Section 3, we discuss how to test completeness and ambiguity for systems using digits 0 and 1 only. In Section 4 we discuss systems using digits  $-1, 0, 1$  only. In Section 5 we discuss representations for all integers, not just the natural numbers. In Section 6 we discuss an entirely new type of Fibonacci representation based on dictionary order. Finally, in Section 7 we describe a few of the new Fibonacci representations we found through exhaustive search of small automata.

## 2 The decision procedure and Walnut

We assume the reader is familiar with the basics of automata theory as discussed, for example, in [10].

The following particular case of a theorem of Büchi [6] (as later corrected by Bruyère et al. [5] is our principal tool in the paper.

**Theorem 1.** *There is a decision procedure that, given a first-order logical formula  $F$  involving natural numbers, comparisons, automata, and addition, and no free variables, will decide the truth or falsity of  $F$ . Furthermore, if  $F$  has free variables, the procedure constructs a DFA accepting those values of the free variables (in Fibonacci representation) that make  $F$  evaluate to TRUE.*

For more information about the specific case of the decision procedure for Fibonacci representation, see [14].

We should explain how automata can process pairs, triples, and generally  $k$ -tuples of inputs. This is done by replacing the input alphabet  $\Sigma$  with the alphabet  $\overbrace{\Sigma \times \Sigma \times \cdots \times \Sigma}^{k \text{ times}}$ . In other words, inputs are  $k$ -tuples of alphabet symbols. The  $i$ 'th input then corresponds to the concatenation of the  $i$ 'th components of all the  $k$ -tuples. Of course, this means that all  $k$  inputs have to have the same length; this is achieved by padding shorter inputs, if necessary, with leading zeros.

The decision procedure of Theorem 1 has been implemented in free software called `Walnut`, originally created by Hamoon Mousavi [13]; also see the book [19]. We recall some of the basics of `Walnut` syntax:

- `eval` evaluates a formula with no free variables and returns TRUE or FALSE; `def` defines an automaton for future use; `reg` defines a regular expression.
- In a regular expression, the period is an abbreviation for the entire alphabet.

- & is logical AND, | is logical OR, => is logical implication, <=> is logical IFF, ~ denotes logical NOT.
- A denotes  $\forall$  (for all); E denotes  $\exists$  (there exists).
- ?msd\_fib tells Walnut to evaluate an arithmetic expression using Fibonacci representation.

We use Walnut to do the computations needed to verify that a given system is complete and unambiguous. For much more about Walnut, including a link to download it, visit

<https://cs.uwaterloo.ca/~shallit/walnut.html> .

### 3 Representation of natural numbers using digits 0 and 1 only

In this section we consider representations of the natural numbers by Fibonacci numbers using digits 0 and 1 only.

The first step is to find an automaton that can convert from an arbitrary Fibonacci representation to the greedy or Zeckendorf representation. To do this we use the following simple observation:

**Proposition 2.** *We can convert a binary string  $x$  to a Zeckendorf representation  $y$  for the same number using the following algorithm: first append a 0 on the front, if necessary. Then scan the string from left to right, replacing each occurrence of “011” successively with “100”.*

*Proof.* Clearly each such replacement does not change the value of  $[x]_F$ . The algorithm terminates because each replacement lowers the total number of 1’s by 1. Finally, the algorithm clearly cannot result in two consecutive 1’s, because it introduces two consecutive 0’s, only the second of which can later change to a 1. □

We can implement this idea as a DFA  $C$  that takes two inputs in parallel,  $x$  and  $y$ , and accepts if and only if both  $[x]_F = [y]_F$  and  $y$  is a valid Zeckendorf representation; that is, it contains no two consecutive 1’s. It suffices to keep track of  $[x']_F - [y']_F$  for the prefix  $x'$  of  $x$  seen so far, and similarly for the prefix  $y'$  of  $y$  seen so far. Note that we assume that  $x$  and  $y$  have the same length, with the shorter of the two prefixed by leading zeros, if necessary. We can think of this as a “converter” or “normalizer” that allows us to turn arbitrary Fibonacci representations into Zeckendorf representations. It is depicted in Figure 1. This automaton was given by Berstel [2] in a slightly different form. Also see [18].

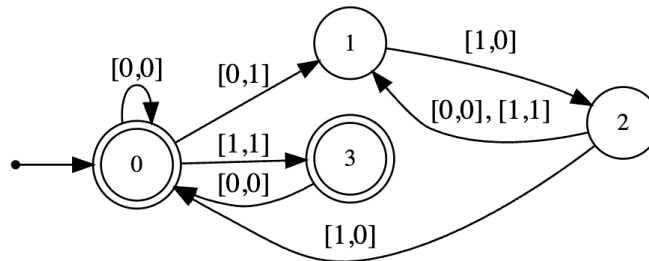


Figure 1: DFA  $C$  for conversion to the Zeckendorf representation.

As an example, consider the input  $[0, 1][1, 0][1, 0][1, 1][0, 0]$  to  $C$ , whose first components spell out  $x = 01110$  and whose second components spell out  $y = 10010$ . Starting in state 0, the automaton visits, successively, states 1, 2, 0, 3, 0, and hence accepts—as it should, since  $[x]_F = [y]_F$ .

We now state one of our main results.

**Theorem 3.** *There is an algorithm that, given rules that specify which representations are valid (in the form of a regular language  $L$  of all valid representations), will decide if the corresponding numeration system based on the Fibonacci numbers is complete and unambiguous for  $\mathbb{N}$ .*

*Proof.* Using Theorem 1, it suffices to express the properties of completeness and unambiguity as a first-order logic formula  $F$ . Once this is done, the decision algorithm can determine if  $F$  is true or false.

Completeness says every integer has a representation in  $L$ . We can express this as follows:

$$\forall n \exists x x \in L \wedge [x]_F = n, \quad (2)$$

Unambiguity says that no integer has two distinct representations in  $L$ . We can express this as follows:

$$\neg \exists x, y \in L (\neg \text{equal}(x, y)) \wedge [x]_F = [y]_F. \quad (3)$$

Here equal means that  $x$  and  $y$  are the same, up to leading zeros.  $\square$

Furthermore, if  $L$  is a regular language that provides a system that is complete, we can find a representation in  $L$  for  $n$  efficiently. The first step is to represent  $n$  in Fibonacci representation, say using the greedy algorithm. Construct a new automaton from `fcanon` by using two intersections. The first intersection is with an automaton with a first component that belongs to  $L$ , while the second component is arbitrary. The second intersection is with an automaton where the first component is arbitrary, and the second is of the form  $0^*(n)_F$ . This gives a new automaton of  $O(\log n)$  states, and it now suffices to find any accepting path (a path from the initial state to the final state). This can be done in linear time in the number of states using depth-first or breadth-first search. This gives us an  $O(\log n)$  algorithm to find a representation. Thus we have proved:

**Theorem 4.** *Suppose  $L$  is a regular language. If  $L$  is complete, we can find a representation for an integer  $n$  in  $O(\log n)$  time.*

*Remark 5.* Here we use the convention of the so-called “word RAM” model, where we assume that  $n$  fits in a single machine word, or more generally that we can perform basic operations on integers with  $O(\log n)$  bits in unit time.

All this can be carried out mechanically with `Walnut`. Here all we have to do is define the language  $L$  of valid representations (say, with a regular expression) and type in the `Walnut` commands corresponding to the two logical assertions (2) and (3). We illustrate this with two examples.

The first is the lazy representation mentioned previously, and discussed first by Brown [3]. The first step is to give a regular expression defining a valid representation in Brown’s system:

```
reg lazyExclude {0,1} "0*1(0|1)*00(0|1)*":
def lazy "~$lazyExclude(s)":
```

This gives a 4-state automaton testing the lazy criterion that is depicted in Figure 2.

We test the completeness and unambiguity for Brown’s system as follows.

```
reg equal {0,1} {0,1} "([0,0] | [1,1])*":
eval brown1 "?msd_fib An Es $fcanon(s,n) & $lazy(s)":
eval brown2 "?msd_fib ~En,s,t $lazy(s) & $lazy(t) & (~$equal(s,t))
& $fcanon(s,n) & $fcanon(t,n) ":
```

Both return TRUE. Given these results, we have now proven that the lazy representation is complete and unambiguous.

For a second example, see the Appendix.

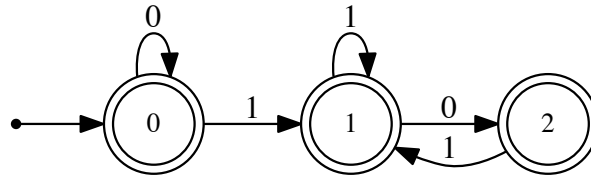


Figure 2: DFA for Brown’s lazy representation.

### 4 Representation using digits $-1, 0,$ and $1$

We now turn to representations using digits  $-1, 0,$  and  $1$  in the Fibonacci system.

Recently, Hajnal [9] described three Fibonacci representations using Eq. (1) to associate a string  $x = e_1 e_{l-1} \cdots e_2 \in \{-1, 0, 1\}^*$  with a natural number  $n$ : *alternating, even, and odd*. Using induction and a case-based argument, he proved that each of these three representations is complete and unambiguous.

Using automata, we can replace his rather long arguments with our general approach. We first describe each of his systems, and show that the set of valid representations for all natural numbers is a regular language.

The alternating representation requires a representation to fulfill four conditions:

1. the most significant nonzero term is positive,
2. two adjacent nonzero terms cannot be of the same sign,
3. two adjacent nonzero terms have at least one zero in between, and
4. if there are two or more nonzero terms, then there has to be at least two zeros between the last and the second last nonzero terms.

We denote a number  $n$  in this representation as  $[n]_A$ . For example,  $[9]_A = 10\bar{1}001$ , where  $\bar{1}$  is used for  $-1$ .

For the alternating representation, we can use the following Walnut code:

```

reg altInclude1 {-1,0,1} "(0*|0*1.*)":
reg altExclude1 {-1,0,1} ".*(10*1|[-1]0*[-1]).*":
reg altExclude2 {-1,0,1} ".*(1[-1]|[-1]1).*":
reg altInclude2 {-1,0,1} "(0*|0*10*|.*(100+[-1]|[-1]00+1)0*)":
def alt "~$altExclude1(s) & ~$altExclude2(s) & $altInclude1(s) & $altInclude2(s)":
  
```

The result is an automaton of 12 states that checks whether an input over the alphabet  $\{-1, 0, 1\}$  is alternating, and is illustrated in Figure 3.

The even representation requires three conditions:

1. the most significant nonzero term is positive,
2. only positions indexed with even numbers, such as  $e_2$ , can have nonzero terms, and
3. two adjacent nonzero terms cannot both be  $-1$ .

We denote a number  $n$  in this representation as  $[n]_E$ . For example,  $[14]_E = 10\bar{1}0001$ .

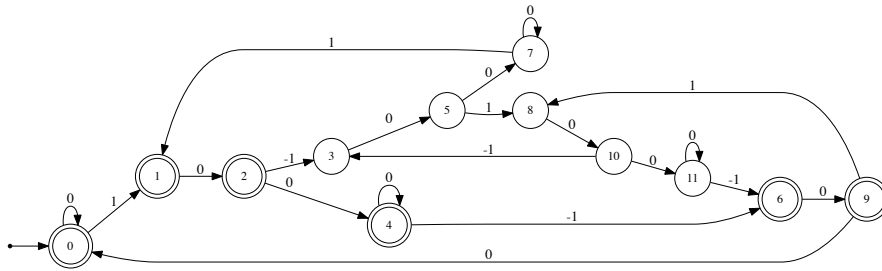


Figure 3: DFA for the alternating condition.

```

reg evenInclude {-1,0,1} "(0*|0*1(0[-1]|01|00)*)":
reg evenExclude {-1,0,1} ".*[-1]0*[-1].*":
def even "$evenInclude(s) & ~$evenExclude(s)":

```

This gives us a 5-state automaton to check the even condition, which is illustrated in Figure 4.

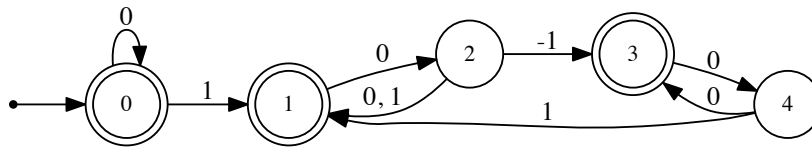


Figure 4: DFA for the even condition.

The odd representation adds an epsilon term to the sum in Eq. (1), therefore associating a string  $e_t e_{t-1} \cdots e_2 \varepsilon$ , where  $\varepsilon \in \{-1, 0\}$ , with a number  $n$ . The odd representation requires the string to meet three conditions:

1. the most significant nonzero term is positive,
2. only positions indexed with odd numbers (such as  $e_3$ ) and the epsilon term are allowed to be nonzero, and
3. two adjacent nonzero terms cannot both be  $-1$ .

We denote a number  $n$  in this representation as  $[n]_O$ . For example,  $[14]_O = 100010\bar{1}$ , where  $\bar{1}$  is used for  $\varepsilon = -1$ .

We express the odd representation conditions in Walnut as follows. Notice we relax the third condition (required in [9]) slightly by limiting its application to only the string  $e_t e_{t-1} \cdots e_2$  without the  $\varepsilon$  term.

```

reg oddInclude {-1,0,1} "(0*|0*10([-1]0|10|00)*)":
reg oddExclude {-1,0,1} ".*[-1]0*[-1].*":
def odd "$oddInclude(s) & ~$oddExclude(s)":

```

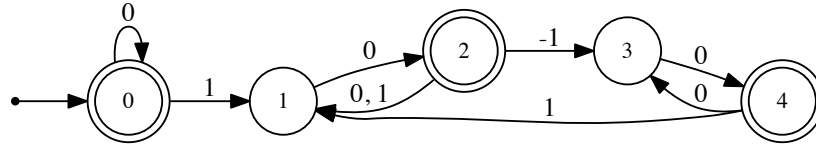


Figure 5: DFA for the odd condition.

This gives us a 5-state automaton to check the odd condition, which is illustrated in Figure 5.

It now remains to use our technique to show that these representations are all complete and unambiguous. In order to do this, we need a “converter” automaton that can compare representations using digits  $-1, 0, 1$  to ordinary Zeckendorf representation. We can construct such an automaton based on `fcanon` as follows. The idea is to use one automaton to “select” the positive digits of a representation, another one to “select” the negative digits, and then do an (implicit) subtraction to obtain the value of the representation.

```

reg posdigits {-1,0,1} {0,1} "([1,1]|[-1,0]|[0,0])*":
reg negdigits {-1,0,1} {0,1} "([-1,1]|[1,0]|[0,0])*":
def fcanon2 "?msd_fib Et,u,w,s $negdigits(x,t) & $posdigits(x,u) &
  $fcanon(t,w) & $fcanon(u,s) & z+w=s":

```

This gives a 24-state automaton `fcanon2`, the analogue of `fcanon`, for doing the conversion.

Let us now check that the alternating representation of Hajnal is both complete and unambiguous.

```

reg same {-1,0,1} {-1,0,1} "([-1,-1]|[0,0]|[1,1])*":
eval altRep1 "?msd_fib An Es $fcanon2(s,n) & $alt(s)":
# evaluates to TRUE, 4 ms
eval altRep2 "?msd_fib ~En,s,t $alt(s) & $alt(t) & (~$same(s,t))
  & $fcanon2(s,n) & $fcanon2(t,n)":
# evaluates to TRUE, 31 ms

```

Similarly, we can check the even and odd representations, as follows:

```

eval evenRep1 "?msd_fib An Es $fcanon2(s,n) & $even(s)":
# evaluates to TRUE, 1 ms
eval evenRep2 "?msd_fib ~En,s,t $even(s) & $even(t) & (~$same(s,t))
  & $fcanon2(s,n) & $fcanon2(t,n)":
# evaluates to TRUE, 4 ms
eval oddRep1 "?msd_fib An (Es $fcanon2(s,n) & $odd(s)) |
  (Et $fcanon2(t,n+1) & $odd(t))":
# evaluates to TRUE, 7 ms
eval oddRep2 "~En,s,t $odd(s) & $odd(t) & (~$same(s,t))
  & $fcanon2(s,n) & $fcanon2(t,n)":
# evaluates to TRUE, 4 ms

```

This completes our proof that all three systems of Hajnal are complete and unambiguous.



*Remark 6.* We noticed, by testing the following, that this representation is also complete if  $\varepsilon \in \{1,0\}$  instead of  $\varepsilon \in \{-1,0\}$  as required in [9].

```
eval oddRep3 "?msd_fib An
  (Es $fcanon2(s,n) & $odd(s)) | (Et $fcanon2(t,n-1) & $odd(t))":
# evaluates to TRUE, 4 ms
```

## 5 Representations for all integers

In this section we investigate two different ways to represent *all* integers (not just the natural numbers) using Fibonacci representations.

Alpert [1] described a *far-difference representation* for Fibonacci numbers that writes *every integer* (not just the natural numbers), with a Fibonacci numeration system using the digits  $-1, 0, 1$ . In Alpert's system, the far-difference representation requires the string to have

1. at least three zeros between any two nonzero terms of the same sign, and
2. at least two zeros between any two nonzero terms of different signs.

We use  $[n]_A$  to denote a natural number in this representation: for example,  $[-38]_A = \bar{1}000\bar{1}001$ . One nice feature of Alpert's system is that it is very easy to negate an integer: all we have to do is change the sign of each digit.<sup>2</sup>

We express the far-difference representation conditions in Walnut as follows.

```
reg exclude1 {-1, 0, 1} ".*([-1][-1]|[-1]0[-1]|[-1]00[-1]|11|101|1001).*":
reg exclude2 {-1, 0, 1} ".*([-1]1|1[-1]|10[-1]|[-1]01).*":
def alpert "~$exclude1(s) & ~$exclude2(s)":
```

This gives a 7-state automaton that checks the Alpert condition, as illustrated in Figure 6.

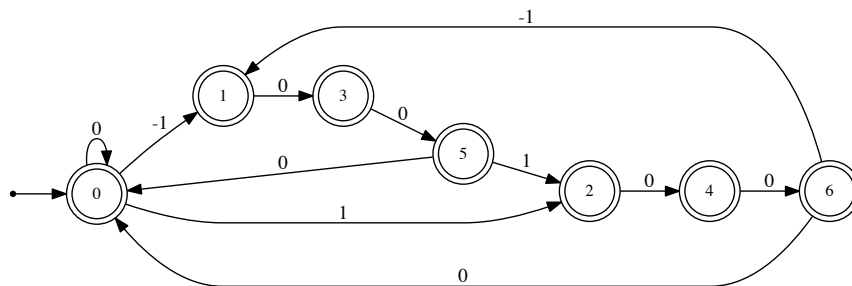


Figure 6: DFA for the Alpert conditions.

To check completeness and ambiguity, we have to check positive and negative integers separately. In addition to `fcanon2`, we need an automaton `fcanon2_neg` that takes a string  $x$  over the alphabet  $\{-1, 0, 1\}$  and a natural number  $n \geq 0$  as input and accepts if  $[x]_F = -n$ .

<sup>2</sup>The three systems proposed by Hajnal also exhibit this property. Therefore, if we exclude the condition stating "the most significant nonzero term is positive" from the three systems, they can be perfect representations for all integers.

```
def fcanon2_neg "?msd_fib Et,u,w,s $negdigits(x,t) & $posdigits(x,u) &
  $fcanon(t,w) & $fcanon(u,s) & z+s=w":
```

We can then prove the completeness and unambiguity of this system as follows.

```
eval farDiff1_pos "?msd_fib An Es $fcanon2(s,n) & $alpert(s)":
eval farDiff1_neg "?msd_fib An Es $fcanon2_neg(s,n) & $alpert(s)":
# both evaluate to TRUE, 3 ms
eval farDiff2_pos "?msd_fib ~En,s,t $alpert(s) & $alpert(t)
  & (~$same(s,t)) & $fcanon2(s,n) & $fcanon2(t,n)":
eval farDiff2_neg "?msd_fib ~En,s,t $alpert(s) & $alpert(t)
  & (~$same(s,t)) & $fcanon2_neg(s,n) & $fcanon2_neg(t,n)":
# both evaluate to TRUE, 9 ms
```

Thus we have easily verified the correctness of Alpert’s conditions.

Bunder [7] invented a different numeration system for all integers, called the negaFibonacci system. In this system, we write integers as a sum of distinct Fibonacci numbers with negative indices, subject to the condition that no two consecutive Fibonacci numbers can be used. Since  $F_{-n} = (-1)^{n+1}F_n$  for  $n \geq 1$ , this is the same as enforcing the requirement in a Fibonacci representation  $a_i F_i + \dots + a_2 F_2 + a_1 F_1$  with digits  $a_i \in \{-1, 0, 1\}$ , (a) only the terms with odd indices are allowed to be positive and only the terms with even indices are allowed to be negative and (b) no two consecutive nonzero digits can appear. We can enforce this condition as follows:

```
reg bunder1 {-1,0,1} ".*1(..)*":
reg bunder2 {-1,0,1} ".*[-1](..)*":
reg bunder3 {-1,0,1} ".*((1[-1])|([-1]1)).*":
def bunder "~$bunder1(x) & ~$bunder2(x) & ~$bunder3(x)":
```

which gives the automaton in Figure 7. We can then check completeness and unambiguity much as

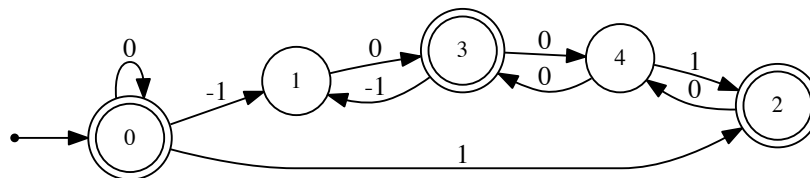


Figure 7: DFA for the Bunder conditions.

we did for Alpert’s system, but there is a new wrinkle: representations have an extra digit at the end, corresponding to the term  $a_1 F_1$ , that must be taken care of. To do this we introduce a “shifter” automaton that shifts a representation to the right, and a “lastbit” that determines if the last bit of a representation is 1 or 0. The shifter is called `rshiftfib` and is displayed in Figure 8.

Then Bunder’s representation can be verified to be complete and unambiguous, as follows:

```
reg lastbit {-1,0,1} {0,1} "([0,0] | [1,0] | [-1,0]) * ([1,1] | [0,0])":
def fcanon3 "?msd_fib Et,u,m $rshiftfib(x,t) &
```

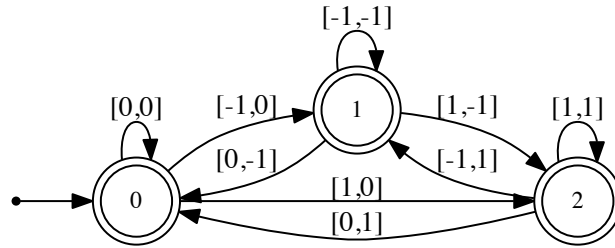


Figure 8: Shifter automaton.

```

$lastbit(x,u) & $fcanon2(t,m) & z=m+u":
def fcanon3_neg "?msd_fib Et,u,m $rshiftfib(x,t) &
  $lastbit(x,u) & $fcanon2_neg(t,m) & z=m-u":
eval bunder1_pos "?msd_fib An Es $fcanon3(s,n) & $bunder(s)":
eval bunder1_neg "?msd_fib An Es $fcanon3_neg(s,n) & $bunder(s)":
# both evaluate to TRUE, 1 ms
eval bunder2_pos "?msd_fib ~En,s,t $bunder(s) & $bunder(t)
  & (~$same(s,t)) & $fcanon3(s,n) & $fcanon3(t,n)":
eval bunder2_neg "?msd_fib ~En,s,t $bunder(s) & $bunder(t)
  & (~$same(s,t)) & $fcanon3_neg(s,n) & $fcanon3_neg(t,n)":
# both evaluate to TRUE, 12 ms

```

## 6 Maximum dictionary order representation

In this section we consider an entirely new Fibonacci representation based on dictionary order. We first introduce how strings are compared in dictionary order. Let  $s = s_1s_2 \cdots s_m$  and  $t = t_1t_2 \cdots t_n$  where  $m \leq n$  be two strings. Let  $i$  such that  $1 \leq i \leq m$  be the first position where  $s_i \neq t_i$ . If  $s_i < t_i$ , then  $s < t$  in dictionary order; otherwise  $s > t$ . For example,  $1011 < 1100$ , but  $1011 > 1001$ . If there is no such position  $i$ , then either  $s = t$  or  $s$  is a proper prefix of  $t$ . In this latter case we say  $s < t$ . For example,  $110 = 110$  and  $110 < 1100$ .

Consider a representation of natural numbers by always choosing the *largest string representation in dictionary order* for every number. Since every number has a Fibonacci-based representation, the representation is complete. Since we choose only one Fibonacci-based representation for each number, the representation is unambiguous. Representations of the first few numbers are given in Table 2.

$n$	1	2	3	4	5	6	7	8	9	10	11
$(n)_D$	1	10	11	101	110	111	1010	1100	1101	1110	1111

Table 2: Representations for the first few numbers.

We now show that

**Theorem 7.** *The set of largest Fibonacci representations in dictionary order forms a regular language.*

*Proof.* The idea is to construct a comparator DFA  $C_D$  that can take two representations in parallel and decide if one is greater than the other, in dictionary order.

In order to take two representations in parallel, they would have to be the same length, and therefore the shorter one would have to be padded with leading zeros to make it the same length as the longer one. In this case, it is not hard to see that no automaton can do the needed comparison.

However, in our case, we can take advantage of the following fact: two Fibonacci representations for the same number cannot be of wildly different lengths.

**Lemma 8.** *The lengths of two Fibonacci-based representation strings for the same natural number differ by one at most (not counting leading zeros).*

*Proof.* Let  $s$  and  $t$  be two Fibonacci representations for a natural number  $m$ . Without loss of generality, assume that  $s$  is longer. Suppose the leading 1 digit of  $s$  corresponds to  $F_i$ . If  $s$  and  $t$  differ in length by more than one, then  $t$  is a sum of some  $F_j$ 's where  $j \leq i - 2$ . Now a classic identity on Fibonacci numbers states that  $\sum_{0 \leq j \leq n} F_j = F_{n+2} - 1$ . Using this relation, we conclude that  $\sum_{j=2}^{i-2} F_j = F_i - 2 < F_i$ . Therefore  $s$  and  $t$  do not represent the same number.  $\square$

Using this fact, it is indeed possible to compare two strings in dictionary order with an automaton. It is shown in Fig. 9 and takes two inputs in parallel,  $s'$  and  $t'$ . Let  $s$  and  $t$  be  $s'$  and  $t'$  without leading

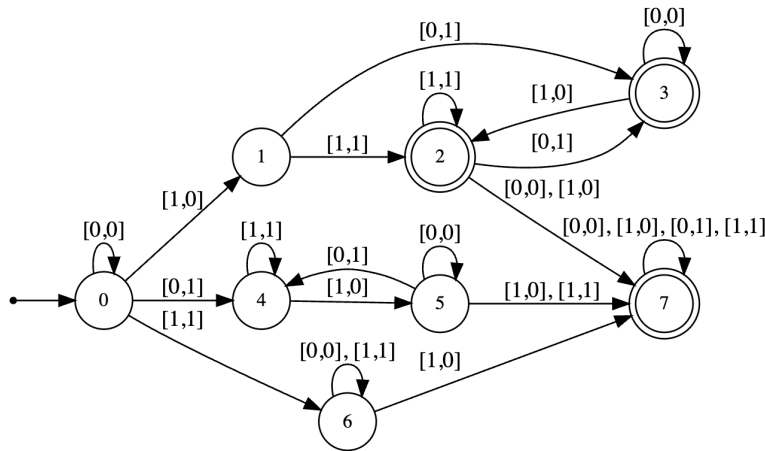


Figure 9: DFA  $C_D$  for comparing strings in dictionary order.

zeros. The DFA  $C_D$  accepts if and only if  $s$  is greater than  $t$  in dictionary order. We have three cases to consider:  $|s| > |t|$ ,  $|s| < |t|$ , and  $|s| = |t|$ . We now discuss how the 8 states of  $C_D$  relate to these 3 cases.

- State 0 is the initial state.
- State 1 is reached if  $|s| > |t|$ ; that is, if  $s'$  starts with 1 and  $t'$  starts with 01.
- State 2 is reached when  $|s| > |t|$ ,  $s$  ends in 1, and based on the inputs so far,  $t$  is a proper prefix of  $s$  therefore  $s > t$ .
- State 3 is reached when  $|s| > |t|$ ,  $s$  ends in 0, and based on the inputs so far,  $t$  is a proper prefix of  $s$  therefore  $s > t$ .
- State 4 is reached when  $|s| < |t|$  and  $t$  ends in 1, and based on the inputs so far,  $s$  is a proper prefix of  $t$  therefore  $s < t$ .

- State 5 is reached when  $|s| < |t|$  and  $t$  ends in 0, and based on the inputs so far,  $s$  is a proper prefix of  $t$  therefore  $s < t$ .
- State 6 is reached when  $|s| = |t|$  and, based on the inputs so far, we have  $s = t$ .
- State 7 is one of the accepting states. It is reached when we can identify a position  $i$  such that  $s_i > t_i$ . Additional symbols read, starting from this state, cannot change the comparison result.

It is now easy to verify that the transitions maintain the invariants corresponding to each state, and we leave this to the reader.  $\square$

Using the comparator automaton, we can build a DFA  $D$  that finds the maximum dictionary order representation for each natural number. The automaton  $D$  takes two inputs in parallel: a number  $n$  in Zeckendorf representation and a string  $s \in \{0, 1\}^*$ ; and it only accepts if, out of all Fibonacci-based representations of  $n$ , the string  $s$  is the greatest based on dictionary order. We implement  $D$  in Walnut as follows.

```
def dictOrder "$fcanon(s,n) & (At $fcanon(t,n) => ($dGreater(s,t)|$equal(s,t)))":
```

Here `dictOrder` implements the automaton  $D$ ; `fcanon`, the automaton  $C$ ; and `dGreater`, the automaton  $C_D$ . The resulting automaton has 7 states and is depicted in Figure 10.

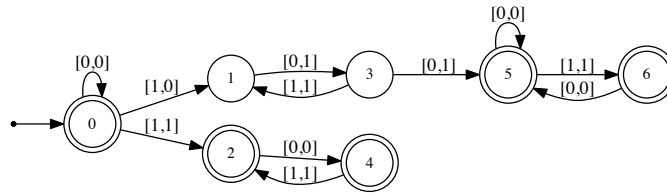


Figure 10: DFA  $D$  for converting to dictionary order representation.

## 7 Finding new perfect systems of small complexity via exhaustive search

We see that a Fibonacci-based representation of natural numbers can be represented by a language over the binary alphabet  $\{0, 1\}$ . If the language is regular, we can express it with a DFA and test its completeness and unambiguity in Walnut. For example, the Zeckendorf representation can be expressed as a 3-state DFA and the Brown one, a 4-state DFA. Therefore we were curious about whether there exist other DFAs with a small number of states that can qualify as complete and unambiguous representations. We conducted an exhaustive search to find such automata and found a surprising number of them. If we allow up to 7 states, we found more than 28 new complete and unambiguous representations.<sup>3</sup> We present two interesting examples out of the seven new 6-state representations.

**Theorem 9.** *Let  $L = 0^*(\varepsilon|1|10(\varepsilon|0|1)1^*(01^+)^*(\varepsilon|0))$ . Then  $L$  is complete and unambiguous.*

*Proof.* We use the following Walnut code:

<sup>3</sup>There could be more as the heuristics we used to trim our search tree can sometimes exclude eligible representations if, for two numbers  $m, n$  where  $m < n$ , the representation of  $m$  is longer than that of  $n$ .

```
reg one0sq {0,1} "0*(()|1|10(()|0|1)1*(01+)*(()|0))":
eval one0sqTestC "?msd_fib An Ex $one0sq(x) & $fcanon(x,n)":
eval one0sqTestU "?msd_fib ~En,x,y $one0sq(x) & $one0sq(y)
& (~$equal(x,y)) & $fcanon(x,n) & $fcanon(y,n)":
```

Both returned TRUE. Here one0sq tests membership in  $L$ . □

Notice this representation allows 100 at the very beginning but no other consecutive 0's are allowed. This restriction on 00 blocks is very similar to Brown's. In fact, Brown's can be expressed, in the form of a regular expression, as

$$0^*(\epsilon|11^*(01^+)^*(\epsilon|0)) = 0^*(\epsilon|1|10(\epsilon|0|1)1^*(01^+)^*(\epsilon|0)).$$

We can imagine that a new representation could be generated for allowing a block of 00 after the second 1, or the third, or after both the first and third 1, or the first and fourth, etc. This offers another construction of infinitely many perfect representations.

**Theorem 10.** *Let  $L$  be the language accepted by the DFA  $Z$ . Then  $L$  is complete and unambiguous.*

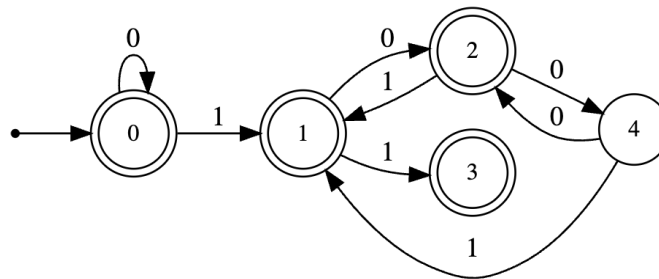


Figure 11: The DFA  $Z$ .

*Proof.* We use the following Walnut code:

```
eval azTestC "?msd_fib An Ex $az(x) & $fcanon(x,n)":
eval azTestU "?msd_fib ~En,x,y $az(x) & $az(y) & (~$equal(x,y))
& $fcanon(x,n) & $fcanon(y,n)":
```

Both returned TRUE. Here az tests membership in  $L$ . □

The strings in  $L$  can end with a single 1 or the block 11 or an odd number of 0's, but not an even number of 0's. Additionally, the strings cannot contain the block "11" anywhere but the end. This restriction on "11" is reminiscent of the Zeckendorf representation.

## 8 Final remarks

The ideas in this paper can be extended in many different ways. For example, we could consider representations in terms of Fibonacci numbers of both positive and negative index with various constraints [16], or representations in terms of sums of the Lucas numbers [4], or other linear recurrences, such as the Pell numbers [11] or Tribonacci numbers [8]. The automaton-based approach can be used in all of these cases.

## References

- [1] H. Alpert (2009): *Differences of multiple Fibonacci numbers*. *INTEGERS* 9, doi:10.1515/INTEG.2009.061. Paper #A57.
- [2] J. Berstel (2001): *An exercise on Fibonacci representations*. *RAIRO Inform. Théor. App.* 35, pp. 491–498, doi:10.1051/ita:2001127.
- [3] J. L. Brown, Jr. (1965): *A new characterization of the Fibonacci numbers*. *Fibonacci Quart.* 3(1), pp. 1–8.
- [4] J. L. Brown, Jr. (1969): *Unique representation of integers as sums of distinct Lucas numbers*. *Fibonacci Quart.* 7, pp. 243–252.
- [5] V. Bruyère, G. Hansel, C. Michaux & R. Villemaire (1994): *Logic and  $p$ -recognizable sets of integers*. *Bull. Belgian Math. Soc.* 1, pp. 191–238, doi:10.36045/bbms/1103408547. Corrigendum, *Bull. Belg. Math. Soc.* 1 (1994), p. 577.
- [6] J. R. Büchi (1960): *Weak second-order arithmetic and finite automata*. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* 6, pp. 66–92, doi:10.1002/malq.19600060105. Reprinted in S. Mac Lane and D. Siefkes, eds., *The Collected Works of J. Richard Büchi*, Springer-Verlag, 1990, pp. 398–424.
- [7] M. W. Bunder (1992): *Zeckendorf representations using negative Fibonacci numbers*. *Fibonacci Quart.* 30, pp. 111–115.
- [8] L. Carlitz, R. Scoville & V.E. Hoggatt, Jr. (1972): *Fibonacci representations of higher order*. *Fibonacci Quart.* 10, pp. 43–69, 94.
- [9] P. Hajnal (2023): *A short note on numeration systems with negative digits allowed*. *Bull. Inst. Combin. Appl.* 97, pp. 54–66.
- [10] J. E. Hopcroft & J. D. Ullman (1979): *Introduction to Automata Theory, Languages, and Computation*. Addison-Wesley.
- [11] A. F. Horadam (1993): *Zeckendorf representations of positive and negative integers by Pell numbers*. In G. E. Bergum, A. N. Philippou & A. F. Horadam, editors: *Applications of Fibonacci Numbers*, 5, Kluwer, pp. 305–316, doi:10.1007/978-94-011-2058-6\_29.
- [12] C. G. Lekkerkerker (1952): *Voorstelling van natuurlijke getallen door een som van getallen van Fibonacci*. *Simon Stevin* 29, pp. 190–195.
- [13] H. Mousavi (2016): *Automatic theorem proving in Walnut*. Arxiv preprint arXiv:1603.06017 [cs.FL], available at <http://arxiv.org/abs/1603.06017>.
- [14] H. Mousavi, L. Schaeffer & J. Shallit (2016): *Decision Algorithms for Fibonacci-Automatic Words, I: Basic Results*. *RAIRO Inform. Théor. App.* 50, pp. 39–66, doi:10.1051/ita/2016010.
- [15] A. Ostrowski (1922): *Bemerkungen zur Theorie der Diophantischen Approximationen*. *Abh. Math. Sem. Hamburg* 1, pp. 77–98, 250–251, doi:10.1007/BF02940595. Reprinted in *Collected Mathematical Papers*, Vol. 3, pp. 57–80.
- [16] H. Park, B. Cho, D. Cho, Y. D. Cho & J. Park (2020): *Representations of integers as sums of Fibonacci numbers*. *Symmetry* 12(10), doi:10.3390/sym12101625. Paper 1625.
- [17] N. Robbins (1996): *Fibonacci partitions*. *Fibonacci Quart.* 34, pp. 306–313.
- [18] J. Shallit (2021): *Robbins and Ardila meet Berstel*. *Inform. Process. Lett.* 167, doi:10.1016/j.ipl.2020.106081. Paper 106081.
- [19] J. Shallit (2022): *The Logical Approach to Automatic Sequences: Exploring Combinatorics on Words with Walnut*. *London Math. Soc. Lecture Notes Series* 482, Cambridge University Press, doi:10.1017/9781108775267.
- [20] E. Zeckendorf (1972): *Représentation des nombres naturels par une somme de nombres de Fibonacci ou de nombres de Lucas*. *Bull. Soc. Roy. Liège* 41, pp. 179–182.