On the powers of the Collatz function

Didier Caucal[1,2] and Chloé Rispal[2]

[1] French National Centre for Scientific Research
[2] Computer Laboratory Gaspard Monge, University Gustave Eiffel, Paris, France
didier.caucal@univ-eiffel.fr chloe.rispal@univ-eiffel.fr

Abstract. For all natural numbers a, b and $d > 0$, we consider the function $f_{a,b,d}$ which associates n/d with any integer n when it is a multiple of d, and $an + b$ otherwise; in particular $f_{3,1,2}$ is the Collatz function. To realize these functions by transducers (automata labelled by pairs of words), the coding in reverse base d is generally used. For the Collatz function, it gives a simple 5-state transducer but it is not suitable for the composition and so far, no one has been able to specify, for all integers p, a generic transducer computing its composition p times. Coding in direct base ad with $b < a$, we realize the functions $f_{a,b,d}$ by synchronous sequential transducers. This particular form makes explicit the composition of such a transducer p times to compute $f^p_{a,b,d}$ in terms of p and a, b, d. We even give an explicit construction of an infinite transducer realizing the closure under composition of $f_{a,b,d}$.

1 Introduction

Many functions on integers have been described by automata (transducers) as word functions using an integer base. In general, the properties of sequences produced by transducers are studied [2] but, in this work, we adress mainly properties of the realized functions themselves.
In this paper, we are interested in the family of functions $f_{a,b} : \mathbb{N} \longrightarrow \mathbb{N}$ defined for all natural numbers a, b and any integer $n \geq 0$ by

$$f_{a,b}(n) \;=\; \begin{cases} \frac{n}{2} & \text{if} \quad n \text{ is even,} \\ an + b & \text{otherwise.} \end{cases}$$

In particular, $f_{3,1}$ is the Collatz function [5]. The Collatz conjecture states that, for any integer n, there exists $p \geq 0$ such that the composition of the Collatz function p times applied to n equals 1: $f^p_{3,1}(n) = 1$. This conjecture remains open despite recent progress [12]. A new conjecture on this function has been proposed [3]. In this paper, our aim is to give an explicit deterministic transducer realizing the p-th power $f^p_{a,b}$ in terms of p.

Basic arithmetic operations have already been described by transducers [7]. To realize the function $f_{a,b}$, the transducer must compute the operations of division by 2, multiplication by a and addition of b. A first natural approach is to take the base 2 with the least significant digit to the left to see right away if the input is even in which case the first digit is 0 which is removed at the output, and if the input is odd, we realize multiplication by a and addition of b

from the left. Thus, the Collatz function can be realized by a 5-state sequential transducer. Introduced by Ginsburg, sequential transducers compute functions deterministically. Their transitions are labelled by a letter in input and a word in ouput with the input-determinism condition: no two transitions of the same input letter from the same source. Furthermore, there is only one initial state and each final state is associated with an output word [8]. The sequential transducer in reverse base 2 for the Collatz function can be composed a few times to realize $f^2_{3,1}$, $f^3_{3,1}$ and so on, but so far, no one has been able to specify, for all integers p, a generic transducer computing its composition p times.

To solve this problem, we chose the direct base $2a$. When $b < a$, which is the case of the Collatz function, we obtain a 2-state deterministic sequential transducer realizing $f_{a,b}$. This new transducer, which in addition is letter-to-letter, can be composed to get an explicit transducer realizing $f^p_{a,b}$ in terms of p. The difficulty in defining its terminal function has been overcome using a key lemma given in [1].

For all natural integers $b < a$ and $d > 0$, we generalize the previous transducers for the functions $f_{a,b,d} : \mathbb{N} \longrightarrow \mathbb{N}$ defined for any integer $n \geq 0$ by

$$f_{a,b,d}(n) \;=\; \begin{cases} \frac{n}{d} & \text{if} \quad n \text{ is a mutiple of } d, \\ an+b & \text{otherwise.} \end{cases}$$

Using the base ad, and for all integer p, we obtain a generic transducer computing $f^p_{a,b,d}$ in terms of p. Finally, for any natural numbers a, b, d with $b < a \neq 1$ and $d \neq 0$, we give an explicit construction of a transducer realizing the closure under composition of $f_{a,b,d}$.

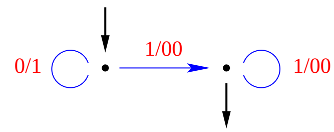
2 Transducers in reverse base 2

In this section, we first recall basic definitions. Then, we look at transducers realizing the functions $f_{a,b}$ using the reverse base 2. Although this approach seems natural, we notice that those transducers are not appropriate for composition.

Let N be a finite alphabet. We denote by N^* the set of words over letters of N, and we write ε for the empty word.

A *transducer* $\mathcal{T} = (T, I, F)$ is a graph defined by a finite edge subset T of $Q{\times}N^*{\times}N^*{\times}Q$, called *transitions*, where Q is a finite set of *states*, plus a set $I \subseteq Q$ of *initial states* and a set $F \subseteq Q$ of *final states*. So a transducer is a finite automaton labelled by pairs of words. Any transition $(p, u, v, q) \in T$ can be also denoted by $p \xrightarrow{u/v}_T q$ or by $p \xrightarrow{u/v} q$ when T is understood; u and v are respectively the *input* and the *output* of the transition.

A *path* $p_0 \xrightarrow{u_1/v_1} p_1 \ldots p_{n-1} \xrightarrow{u_n/v_n} p_n$ with $u = u_1...u_n$ and $v = v_1...v_n$ is labelled by u/v and is denoted by $p_0 \overset{u/v}{\Longrightarrow}_T p_n$. A path is *successful* if it leads from an initial state to a final one. A pair $(u, v) \in N^*{\times}N^*$ is *recognized* by a transducer if there exists a successful path labelled by u/v. The set of recognized pairs is the relation $\langle\mathcal{T}\rangle$ *realized* by $\mathcal{T}$ and called a *rational relation*. For instance, the following transducer:



with a unique initial state on the left and a unique final state on the right realizes the relation $\{ (0^m1^n , 1^m0^{2n}) \mid m \geq 0, n > 0 \}$.
Note that the inverse R^{-1} of a rational relation R is rational, and the image $R(L)$ by R of a regular language L is regular since $R(L) = \pi_2(R \cap L{\times}N^*)$.

To realize the functions $f_{a,b}$ in reverse base 2, we only need synchronized and deterministic transducers.
A transducer $\mathcal{T} = (T, I, F)$ is *synchronized* if T is *letter-to-word* in the following sense: $T \subset Q{\times}N{\times}N^*{\times}Q$ *i.e.* the inputs are only letters.
Given synchronized transducers $\mathcal{T} = (T, I, F)$ and $\mathcal{T}' = (T', I', F')$, their *composition* is the following synchronized transducer:

$$\mathcal{T} \circ \mathcal{T}' = (T \circ T', I{\times}I', F{\times}F')$$

$$\text{where } T \circ T' = \{ (p, p') \xrightarrow{a/v} (q, q') \mid \exists\, u\ (p \xrightarrow{a/u}_T q \wedge p' \overset{u/v}{\Longrightarrow}_{T'} q') \}$$

realizing the composition of the relation realized by $\mathcal{T}$ with that by $\mathcal{T}'$:

$$\langle \mathcal{T} \circ \mathcal{T}' \rangle = \langle \mathcal{T} \rangle \circ \langle \mathcal{T}' \rangle = \{ (u, w) \mid \exists\, v\ (u, v) \in \langle \mathcal{T} \rangle \wedge (v, w) \in \langle \mathcal{T}' \rangle \}.$$

A synchronized transducer $\mathcal{T} = (T, I, F)$ is *deterministic* if it has a unique state *i.e.* $|I| = 1$, and its graph T is *input-deterministic* in the following sense:

$$(p \xrightarrow{a/u} q \wedge p \xrightarrow{a/v} r) \implies (u = v \wedge q = r).$$

A deterministic synchronized transducer realizes a function.
A synchronized transducer $\mathcal{T} = (T, I, F)$ is *complete* if its graph of vertex set Q is *input-complete* in the following sense:

$$\forall\, p \in Q, \forall\, a \in N, \exists\, u \in N^*, \exists\, q \in Q \ \ p \xrightarrow{a/u} q.$$

A deterministic and complete synchronized transducer realizes an application.

Let $\beta > 1$ be an integer and $\widehat{\beta} = \{0, \ldots, \beta - 1\}$ be the alphabet of its *digits*. Any word $u \in \widehat{\beta}^*$ is a (respectively reverse) *representation in base* β of the integer $[u]_\beta$ (respectively ${}_\beta[u]$) defined by $[\varepsilon]_\beta = 0 = {}_\beta[\varepsilon]$ and

$$[c_p \ldots c_0]_\beta = \textstyle\sum_{i=0}^{p} c_i \beta^i = {}_\beta[c_0 \ldots c_p]$$

for any $p \geq 0$ and $c_0, \ldots, c_p \in \widehat{\beta}$; the position of the index β is that of the least significant digit c_0. Representations of integers are extended to relations. A relation $R \subseteq \widehat{\beta}^*{\times}\widehat{\beta}^*$ is a (resp. reverse) representation in base β of the following binary relation $[R]_\beta$ (respectively ${}_\beta[R]$) on $\mathbb{N}$:

$$[R]_\beta = \{ ([u]_\beta, [v]_\beta) \mid (u, v) \in R \} \text{ and } {}_\beta[R] = \{ ({}_\beta[u], {}_\beta[v]) \mid (u, v) \in R \}.$$

The functions $f_{a,b}$ on integers can be seen as relations on words and defined by transducers. First, let us see what transducers can be obtained for the functions $f_{a,b}$ using a coding in reverse base 2. For some simple functions such as $f_{1,1}$, we get a transducer for the composition p times using a shortcut. This is not possible in the general case. Consider the function $f_{1,1} : \mathbb{N} \longrightarrow \mathbb{N}$ defined for any integer $n \geq 0$ by

$$f_{1,1}(n) \;=\; \begin{cases} \frac{n}{2} & \text{if} \quad n \text{ is even,} \\ n+1 & \text{otherwise.} \end{cases}$$

The first natural approach is to take the base 2 with the least significant digit to the left. In reverse base 2, $f_{1,1}$ is represented by the following word function:

$$\begin{array}{rcll} 0u & \longrightarrow & u & \text{for any binary word } u \\ 1u & \longrightarrow & 0(u+1) & \end{array}$$

This word function is realized by the deterministic synchronized 3-state transducer:

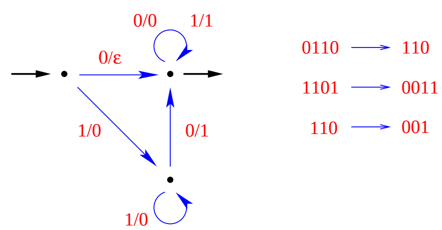


Fig. 1. Transducer realizing $f_{1,1}$ in reverse base 2.

Taking the product p times with itself, we can construct an automaton for $f_{1,1}^p$ having 3^p states but we do not know how to define it in terms of p. To solve this problem, for any natural integers a, b of same parity, we consider the *short-cut* $f'_{a,b}$ defined for any natural number n by

$$f'_{a,b}(n) \;=\; \begin{cases} \frac{n}{2} & \text{if} \quad n \text{ is even,} \\ \frac{an+b}{2} & \text{otherwise.} \end{cases}$$

In the particular case of $f'_{1,1}$, we get the property below.

Lemma 1. *For all* $n, p \geq 0$, $f'^{\,p}_{1,1}(n) \;=\; \lceil \frac{n}{2^p} \rceil$.

Proof. By induction on $p \geq 0$.
$p = 0$: For all $n \geq 0$, $f'^{\,0}_{1,1}(n) = n = \lceil \frac{n}{2^0} \rceil$.
$p > 0$: If n is even then $f'^{\,p}_{1,1}(n) = f'^{\,p-1}_{1,1}(\frac{n}{2}) = \lceil \frac{n}{2^p} \rceil$ by induction hypothesis. Otherwise n is odd *i.e.* $n = 2^p k + r$ for some $k \geq 0$ and $0 < r < 2^p$ odd. So $f'^{\,p}_{1,1}(n) \;=\; f'^{\,p-1}_{1,1}(\frac{n+1}{2}) \;=\; \lceil \frac{n+1}{2^p} \rceil \;=\; k + \lceil \frac{r+1}{2^p} \rceil \;=\; k+1 \;=\; k + \lceil \frac{r}{2^p} \rceil \;=\; \lceil \frac{n}{2^p} \rceil$. ◀

So for any $n, p \geq 0$,

$$f'^{\,p}_{1,1}(n) \;=\; \begin{cases} \frac{n}{2^p} & \text{if} \quad n \text{ is a multiple of } 2^p, \\ \lfloor \frac{n}{2^p} \rfloor + 1 & \text{otherwise.} \end{cases}$$

In reverse base 2, we get the function

$$\begin{array}{rcll} 0^p u & \longrightarrow & u & \text{for any } u \in \{0,1\}^* \\ vu & \longrightarrow & u+1 & \text{for } |v| = p \text{ and } v \neq 0^p \end{array}$$

which is realized by the following transducer:

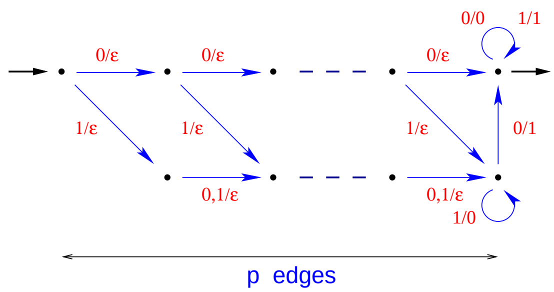


Fig. 2. Transducer realizing $f'^{\,p}_{1,1}$ in reverse base 2.

In this case, the shortcut breaks the exponential number 3^p of states into the linear number $2p+1$. Let us try to do this for the *Collatz function*:

$$f_{3,1}(n) \;=\; \begin{cases} \frac{n}{2} & \text{if} \quad n \text{ is even,} \\ 3n+1 & \text{otherwise.} \end{cases}$$

Using the reverse base 2, we get the following deterministic and complete synchronized transducer $\mathcal{C}$ among others [6, 10].

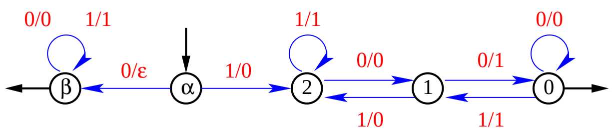


Fig. 3. A transducer realizing the Collatz function in reverse base 2.

The states 0, 1, 2 manage the carry of the multiplication by 3.

From such a transducer, it follows a property to describe the behaviour of the Collatz function by automata [9].

Precisely, and for any integer n, we denote by

$$\lambda(n) \;=\; |\{\; p \mid f^p_{3,1}(n) \text{ odd and } f^q_{3,1}(n) \neq 1 \text{ for } 0 \leq q < p \;\}|$$

the total of odd numbers (rises) of the orbit of $f_{3,1}$ from n until 1 is possibly reached. In particular $\lambda^{-1}(0)$ is the set of powers of 2. For any $i \geq 0$, let

$$L_i \;=\; \{\; u \in \{0,1\}^*1 \cup \{\varepsilon\} \mid \lambda({}_2[u]) \;=\; i \;\}$$

be the set of reverse binary representations of $\lambda^{-1}(i)$ with non null rightmost digit. In particular

$$L_0 \;=\; 0^*1 \cup \{\varepsilon\} \quad \text{is a regular language.}$$

This regularity remains true for all languages.

Lemma 2. *For any $i \geq 0$, L_i is a regular language.*

Proof.
Let $f : \{0,1,2\}^* \longrightarrow \{0,1,2\}^*$ be the application realized by the transducer of

Figure 3 :

$$f_{3,1}({}_2[u]) \;=\; {}_2[f(u)] \;\text{ for any }\; u \in \{0,1,2\}^*.$$

The proof is done by induction on $i \geq 0$. We have

$$\begin{aligned}
& u \in L_{i+1} \\
\Longleftrightarrow\; & u \in \{0,1\}^*1 \;\text{ and }\; \lambda({}_2[u]) \;=\; i+1 \\
\Longleftrightarrow\; & u \in \{0,1\}^*1 \;\text{ and }\; \exists\, p,q \geq 0 \;({}_2[u] \;=\; 2^p q \,\wedge\, q \text{ odd } \wedge\, \lambda(f_{3,1}(q)) \;=\; i) \\
\Longleftrightarrow\; & \exists\, v \in \{1\} \,\cup\, 1\{0,1\}^*1 \;\big(u \in 0^*v \,\wedge\, \lambda({}_2[f(v)]) \;=\; i\big) \\
\Longleftrightarrow\; & \exists\, v \in (1^*0)^*1^+ \;\big(u \in 0^*v \,\wedge\, f(v) \in L_i\, 0^*\big) \\
\Longleftrightarrow\; & u \in 0^*\big(f^{-1}(L_i\, 0^*) \,\cap\, (1^*0)^*1^+\big).
\end{aligned}$$

For L_i regular, $L_{i+1} \;=\; 0^*\big(f^{-1}(L_i\, 0^*) \,\cap\, (1^*0)^*1^+\big)$ is a regular language. ◀

By Lemma 2, the reverse $\{\; u \in 1\{0,1\}^* \cup \{\varepsilon\} \mid \lambda([u]_2) \;=\; i \;\}$ of L_i is a regular language [9]. For all i, a regular expression defining L_i can be constructed in exponential in i [11].

Let us return to the transducer $\mathcal{C}$ of Figure 3. Its composition twice $\mathcal{C} \circ \mathcal{C}$ is

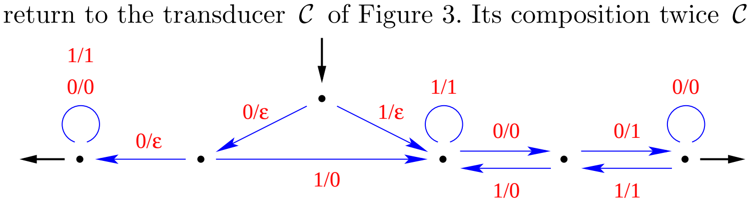


Fig. 4. A transducer realizing the power of 2 of the Collatz function.

where the states $(2,\beta)$ and $(\beta,2)$ are identified since they are *equivalent*: $\mathcal{C}$ realizes the same function from each one. This last transducer realizes

$$f_{3,1}^2(n) \;=\; \begin{cases} \frac{n}{4} & \text{if } \; n \in 4\,\mathbb{N} \\ \frac{3n+2}{2} & \text{if } \; n \in 4\,\mathbb{N} \,+\, 2 \\ \frac{3n+1}{2} & \text{if } \; n \text{ is odd.} \end{cases}$$

By identifying equivalent vertices of the composition 3 times of $\mathcal{C}$, we get the following transducer:

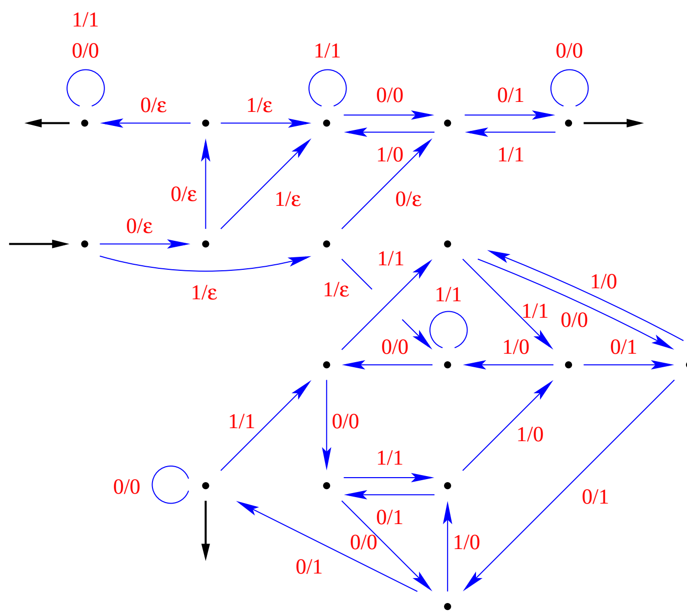


Fig. 5. A transducer realizing the power of 3 of the Collatz function.

The heaviness of this transducer comes from the fact that it performs each of the cases of the function

$$f_{3,1}^3(n) \;=\; \begin{cases} \frac{n}{8} & \text{if } \; n \in 8\,\mathbb{N} \\ \frac{3n+1}{4} & \text{if } \; n \in 8\,\mathbb{N}+1 \;\text{ or }\; n \in 8\,\mathbb{N}+5 \\ \frac{3n+2}{4} & \text{if } \; n \in 8\,\mathbb{N}+2 \;\text{ or }\; n \in 8\,\mathbb{N}+6 \\ \frac{9n+5}{2} & \text{if } \; n \in 8\,\mathbb{N}+3 \;\text{ or }\; n \in 8\,\mathbb{N}+7 \\ \frac{3n+4}{4} & \text{if } \; n \in 8\,\mathbb{N}+4 \end{cases}$$

Thus and contrary to the function $f_{1,1}$, we cannot get a transducer in reverse base 2 for $f_{3,1}^p$ in terms of p, not even for its shortcut $f'^p_{3,1}$.

3 Transducers for the Euclidean division

Before giving other transducers to realize the functions $f_{a,b}$, we recall how to realize an Euclidean division by a transducer.

We only need transducers that are both synchronous and deterministic. A transducer $\mathcal{T} = (T, I, F)$ is *synchronous* if T is *letter-to-letter* in the following sense: $T \subseteq Q{\times}N{\times}N{\times}Q$ *i.e.* the inputs and outputs are only letters. Note that for synchronous transducers $\mathcal{T} = (T, I, F)$ and $\mathcal{T}' = (T', I', F')$, the composition $T \circ T'$ can be expressed more simply as follows:

$$T \circ T' \ = \ \{ \ (p, p') \xrightarrow{a/c} (q, q') \mid \exists\, b \ (p \xrightarrow{a/b}_T q \wedge p' \xrightarrow{b/c}_{T'} q') \ \}.$$

We realize the division by $d > 0$ in base $a > 1$ with remainder $r < d$ by the following standard deterministic synchronous transducer:

$$/_{a,d,r} \ = \ (\widehat{d}\,,\ :_{a,d}\,, \{0\}, \{r\})$$

where

$$i \xrightarrow{b/c}_{:_{a,d}} j \quad \text{if} \quad ia + b \ = \ cd + j \ \text{ for all } \ i, j \in \widehat{d} \ \text{ and } \ b, c \in \widehat{a}.$$

The division $:_{a,d}$ by d in base a is illustrated below.

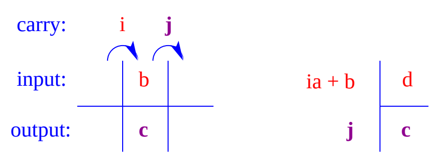


Fig. 6. Division by d in base a for a digit b with a carry i.

This illustration is extended from digits to numbers.

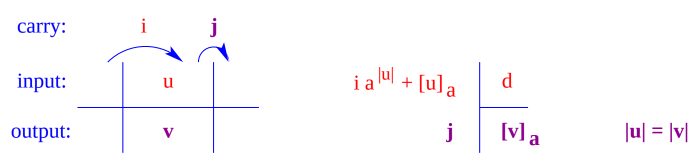


Fig. 7. Division by d in base a for a number $[u]_a$ with a carry i.

We thus extend the transitions of the division to its paths.

Lemma 3. *For all* $i, j \in \widehat{d}$ *and* $u, v \in \widehat{a}^*$, *we have*

$$i \overset{u/v}{\Longrightarrow}_{:_{a,d}} j \quad \Longleftrightarrow \quad i\, a^{|u|} + [u]_a \ = \ [v]_a d + j \ \text{ and } \ |u| = |v|.$$

Proof. Each implication can be checked easily by induction on $|u| \geq 0$.
$\Longrightarrow$: As $:_{a,d}$ is a subset of $\widehat{d} \times \widehat{a} \times \widehat{a} \times \widehat{d}$, $|u| = |v|$.

Let us check the equality by induction on $|u| \geq 0$.
$|u| = 0$: We have $u = \varepsilon = v$ and $i = j$ hence the equality.
Let $i \overset{ub/vc}{\Longrightarrow} j$ with $b, c \in \widehat{a}$ and the implication true for u. There exists k such that $i \overset{u/v}{\Longrightarrow} k \overset{b/c}{\longrightarrow} j$. Thus $i\, a^{|u|} + [u]_a \;=\; [v]_a d + k$ and $ka + b = cd + j$. Hence

$$\begin{aligned} i\, a^{|ub|} + [ub]_a \;&=\; b + (i\, a^{|u|} + [u]_a)a \;=\; b + ([v]_a d + k)a \;=\; [v]_a da + (ka + b) \\ &=\; [v]_a da + cd + j \qquad =\; [vc]_a d + j\,. \end{aligned}$$

$\Longleftarrow$: by induction on $|u| \geq 0$.
$|u| = 0$: We have $u = \varepsilon = v$ and $i = j$ hence $i \overset{u/v}{\Longrightarrow} j$.
Suppose the implication true for $|u|$ and $i\, a^{|ub|} + [ub]_a = [vc]_a d + j$ with $|u| = |v|$ and $0 \leq b, c < a$. So, we have $(i\, a^{|u|} + [u]_a)a + b = [v]_a ad + cd + j$.
By Euclidean division of $cd + j$ by a, we have $cd + j = ka + b'$ with $b' < a$.
As $b < a$, we have $b = b'$ hence $i\, a^{|u|} + [u]_a \;=\; [v]_a d + k$.
As $|u| \;=\; |v|$ and by induction hypothesis, $i \overset{u/v}{\Longrightarrow} k$.
As $cd + j = ka + b$, we get $k \overset{b/c}{\longrightarrow} j$ hence $i \overset{ub/vc}{\Longrightarrow} j$. ◀

Thus $/_{a,d,r}$ realizes the binary relation

$$\{\ (u, v) \mid u, v \in \widehat{a}^* \ \wedge\ |u| = |v| \ \wedge\ [u]_a \;=\; [v]_a d + r\ \}.$$

Here is a representation of $/_{3,2,0}$.

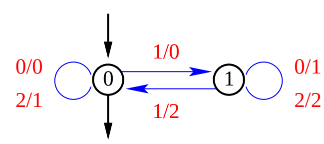


Fig. 8. Division by 2 in base 3 with remainder 0.

For a same base a, the composition $:_{a,d} \circ :_{a,d'}$ is in bijection with the relation $:_{a,dd'}$ by coding any vertex (i, i') where $0 \leq i < d$ and $0 \leq i' < d'$ by the integer ${}_d[(i, i')] \;=\; i + i'd$.

Lemma 4. *For all $a > 1$ and $d, d' > 0$, ${}_d[:_{a,d} \circ :_{a,d'}]$ is equal to $:_{a,dd'}$.*

Proof. For all $0 \leq i, j < d$ and $0 \leq i', j' < d'$, we have

$$\begin{aligned} & (i, i') \overset{b/c}{\longrightarrow}_{:_{a,d} \circ :_{a,d'}} (j, j') \\ \Longleftrightarrow\;& \exists\, 0 \leq e < a \text{ such that } i \overset{b/e}{\longrightarrow}_{:_{a,d}} j \text{ and } i' \overset{e/c}{\Longrightarrow}_{:_{a,d'}} j' \\ \Longleftrightarrow\;& \exists\, 0 \leq e < a \text{ such that } ia + b \;=\; ed + j \text{ and } i'a + e \;=\; cd' + j' \\ \Longleftrightarrow\;& ia + b \;=\; (cd' + j' - i'a)d + j \\ \Longleftrightarrow\;& (i + i'd)a + b \;=\; cdd' + j + j'd \\ \Longleftrightarrow\;& {}_d[(i, i')] \overset{b/c}{\longrightarrow}_{:_{a,dd'}} {}_d[(j, j')]\,. \end{aligned}$$

◀

Let us propose a way to visualize these transducers to highlight basic symmetries. The d integers of the vertex set $\widehat{d} = \{0, \dots, d-1\}$ of $:_{a,d}$ are equidistant on a counterclockwise circle in a way that the diameter between 0 and $d-1$ is horizontal with 0 at the top right. Here is a representation of the transducer in base $a = 2$ for respectively $d = 1, 2, 3, 4$:

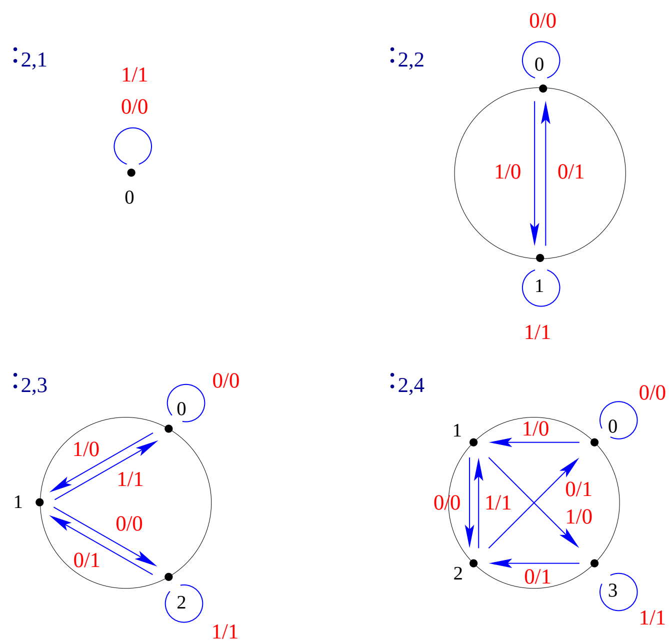


Fig. 9. Visualization of the division.

By associating a color with each digit, any transition $\xrightarrow{b/c}$ is replaced by an unlabeled two-colored arrow: the start of the arrow is with the color of input b, and the end of the arrow is with the color of output c.
In base 3, we color $0, 1, 2$ with blue, green and red respectively. This gives the representation below of division by 8 in base 3.

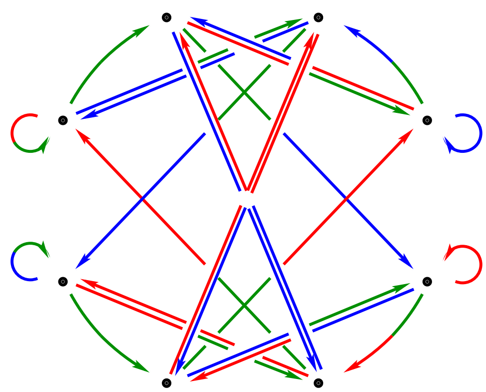


Fig. 10. Division by 8 in base 3.

We can then notice that
- the horizontal diameter is axe of symmetry where we exchange 0 with 2 in input and output, leaving 1 unchanged,
- the vertical diameter is axe of symmetry where we exchange in input 0 with 2, in north output 0 with 1, and in south output 1 with 2,
- the center of the circle is center of symmetry where the input remains the same and for the output, we exchange 0 north with 1 south, 1 north with 2 south, 2 north with 0 south.

Let us point out that we do not know how to characterize powers of 2 in base 3. In particular, there is the Erdös conjecture [4] stating that these powers from 9 have at least a 2 in base 3:

$$u \notin \{0,1\}^* \quad \text{for} \quad [u]_3 = 2^n \quad \text{with} \quad n > 8$$

which translates for the Euclidean division $:_{3,2^n}$ by 2^n in base 3 that the elementary cycle in 0 with output in 0^*1 has its input having at least one 2:

$$u \notin \{0,1\}^* \quad \text{for} \quad 0 \overset{u/0...01}{\Longrightarrow}_{:_{3,2^n}} 0 \quad \text{with} \quad n > 8.$$

Thus, we consider the *composition closure* of the division by d in base a to be

$$:_{a,d}^* = \bigcup_{n \geq 0} :_{a,d}^n$$

the reflexive and transitive closure under composition of $:_{a,d}$.
This infinite relation, of vertex set $\widehat{d}^*$, can be described by the paths of a finite relation.

Let $T \subseteq M \times N \times N \times M$ be a letter-to-letter transition set whose vertices are letters in an alphabet M. The *dual* of T is the transition set $\widetilde{T} \subseteq N \times M \times M \times N$ defined by

$$b \overset{p/q}{\longrightarrow}_{\widetilde{T}} c \quad \Longleftrightarrow \quad p \overset{b/c}{\longrightarrow}_T q \quad \text{for any} \quad p, q \in M \quad \text{and} \quad b, c \in N.$$

Here is a representation of the dual of $:_{3,2}$.

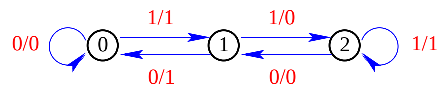


Fig. 11. Dual of the division by 2 in base 3.

The *composition closure* of T is the transition set

$$T^* \;=\; \bigcup_{n\geq 0} T^n \;\subseteq\; M^*{\times}N{\times}N{\times}M^*.$$

This infinite set is described by the paths of the finite set $\widetilde{T}$.

Lemma 5. *For any* $T \subseteq M{\times}N{\times}N{\times}M$ *over alphabets* M *and* N*, we have*

$$x \xrightarrow{b/c}_{T^*} y \quad\Longleftrightarrow\quad b \overset{x/y}{\Longrightarrow}_{\widetilde{T}} c \;\text{ for any } x,y\in M^* \text{ and } b,c\in N.$$

Proof. It suffices to check by induction on $n\geq 0$ that

$$x \xrightarrow{b/c}_{T^n} y \quad\Longleftrightarrow\quad b \overset{x/y}{\Longrightarrow}_{\widetilde{T}} c \;\text{ for any } x,y\in M^n \text{ and } b,c\in N.$$

$n=0$: $x=y=\varepsilon$. Thus $\varepsilon \xrightarrow{b/c}_{T^0} \varepsilon \iff b=c \iff b \overset{\varepsilon/\varepsilon}{\Longrightarrow}_{\widetilde{T}} c$.

$n \Longrightarrow n+1$: For any $x,y\in M^n$ and $p,q\in M$, we have

$$\begin{aligned} xp \xrightarrow{b/c}_{T^{n+1}} yq &\iff \exists\, e\in N \;\; x \xrightarrow{b/e}_{T^n} y \;\text{ and }\; p \xrightarrow{e/c}_{T} q \\ &\iff \exists\, e\in N \;\; b \overset{x/y}{\Longrightarrow}_{\widetilde{T}} e \xrightarrow{p/q}_{\widetilde{T}} c \\ &\iff b \overset{xp/yq}{\Longrightarrow}_{\widetilde{T}} c. \;\blacktriangleleft \end{aligned}$$

We can extend Lemma 5 to the composition closure of the dual.

Lemma 6. *For any* $T \subseteq M{\times}N{\times}N{\times}M$ *over alphabets* M *and* N*, we have*

$$u \overset{x/y}{\Longrightarrow}_{\widetilde{T}^*} v \quad\Longleftrightarrow\quad x \overset{u/v}{\Longrightarrow}_{T^*} y \;\text{ for any } x,y\in M^* \text{ and } u,v\in N^*.$$

Proof. It suffices to check by induction on $n\geq 0$ that

$$u \overset{x/y}{\Longrightarrow}_{\widetilde{T}^n} v \quad\Longleftrightarrow\quad x \overset{u/v}{\Longrightarrow}_{T^*} y \;\text{ for any } x,y\in M^* \text{ and } u,v\in N^n.$$

$n=0$: $u=v=\varepsilon$. Thus $\varepsilon \overset{x/y}{\Longrightarrow}_{\widetilde{T}^0} \varepsilon \iff x=y \iff x \overset{\varepsilon/\varepsilon}{\Longrightarrow}_{T^*} y$.

$n \Longrightarrow n+1$: For any $x,y\in M^*$ and $u,v\in N^n$ and $b,c\in N$, we have

$$\begin{aligned} ub \overset{x/y}{\Longrightarrow}_{\widetilde{T}^{n+1}} vc &\iff \exists\, z\in M^* \;\; u \overset{x/z}{\Longrightarrow}_{\widetilde{T}^n} v \;\text{ and }\; b \overset{z/y}{\Longrightarrow}_{\widetilde{T}} c \\ &\iff \exists\, z\in M^* \;\; x \overset{u/v}{\Longrightarrow}_{T^*} z \;\text{ and }\; z \xrightarrow{b/c}_{T^*} y \quad \text{by Lemma 5} \\ &\iff x \overset{ub/vc}{\Longrightarrow}_{T^*} y. \;\blacktriangleleft \end{aligned}$$

We will now see that the dual of a division is a multiplication.
Let us recall how to realize a multiplication by a transducer.
The set $*_{a,d}$ of transitions for the multiplication by d in reverse base a is defined by

$$i \xrightarrow{b/c}_{*_{a,d}} j \quad\text{if}\quad bd+i \;=\; ja+c \;\text{ for all } i,j\in \widehat{d} \text{ and } b,c\in \widehat{a}$$

which is illustrated below.

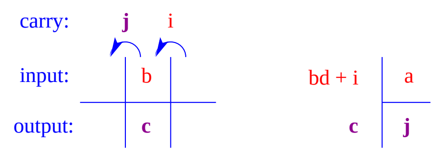


Fig. 12. Multiplication by d in reverse base a for a digit b with a carry i.

Thus, we have

$$i \xrightarrow{b/c}_{*_{a,d}} j \iff j \xrightarrow{c/b}_{:_{a,d}} i \text{ for any } i,j \in \widehat{d} \text{ and } b,c \in \widehat{a}.$$

The paths of $*_{a,d}$ are deduced from Lemma 3.

Corollary 1. *For all $i,j \in \widehat{d}$ and $u,v \in \widehat{a}^*$, we have*

$$i \overset{u/v}{\Longrightarrow}_{*_{a,d}} j \iff j \overset{\widetilde{v}/\widetilde{u}}{\Longrightarrow}_{:_{a,d}} i \iff {}_a[u]\,d + i = {}_a[v] + j\,a^{|v|} \text{ and } |u| = |v|.$$

In particular, we have

$$0 \overset{u/v}{\Longrightarrow}_{*_{a,d}} 0 \iff {}_a[u]\,d = {}_a[v] \text{ and } |u| = |v|.$$

The multiplication by d in reverse base a is then realized by the following synchronous sequential transducer:

$$\times_{a,d} = (\widehat{d}\,,\, *_{a,d}\,, \{0\}, \{0\})$$

From Figure 8, we obtain the following representation of $\times_{3,2}$

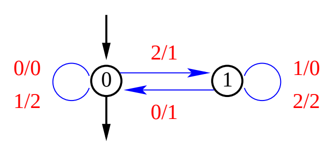


Fig. 13. Multiplication by 2 in reverse base 3.

e.g. this transducer realizes $(02120, 01021)$ which in reverse base 3 gives $(69, 138)$. The dual of the division by d in base a is the multiplication by a in reverse base d.

Lemma 7. *We have $\widetilde{:_{a,d}} = *_{d,a}$ for any $a,d > 1$.*

Proof. For any $b,c \in \widehat{a}$ and $i,j \in \widehat{d}$, we have

$$\begin{aligned} b \xrightarrow{i/j}_{\widetilde{:_{a,d}}} c &\iff i \xrightarrow{b/c}_{:_{a,d}} j \\ &\iff ia + b = cd + j \\ &\iff b \xrightarrow{i/j}_{*_{d,a}} c. \end{aligned}$$ ◀

Thus, Figure 11 is also the transition set of the multiplication by 3 in reverse base 2. We can now realize the Collatz function and its powers with simple transducers.

4 Transducer for $f_{a,b}$ in base $2a$

For any $0 \le b < a$ of same parity, we give a transducer in base a realizing the shortcut $f'_{a,b}$. Since $f_{a,b} = f'_{2a,2b}$, we obtain a transducer in base $2a$ for $f_{a,b}$.

We will realize these functions by deterministic synchronous transducers where each final state is associated with an output word.
A *sequential transducer* [8] is a deterministic synchronized transducer $\mathcal{T} = (T, i, \omega)$ whose the set of final states is extended to a partial *terminal function* $\omega : Q \to N^*$: its domain $\mathrm{dom}(\omega)$ is the set of final states.
We denote by $q \xrightarrow{w}$ when q is a final state such that $\omega(q) = w$.
Such a transducer realizes the binary relation

$$\langle \mathcal{T} \rangle \ = \ \{ \ (u, vw) \mid \exists \ i \in I, \ q \in \mathrm{dom}(\omega) \ (i \overset{u/v}{\Longrightarrow}_T q \xrightarrow{w}) \ \}.$$

For instance, the following synchronous sequential transducer $\mathcal{T}_{2,1}$:

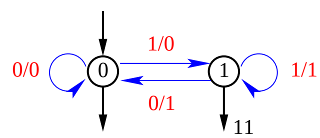


Fig. 14. Transducer realizing $f_{2,1}$ in base 2.

realizes the following word function $\langle \mathcal{T}_{2,1} \rangle$:

$$\begin{aligned} \varepsilon &\longrightarrow \varepsilon \\ u0 &\longrightarrow 0u \quad \text{ for any } u \in \{0,1\}^* \\ u1 &\longrightarrow 0u11 \end{aligned}$$

which is, in direct base 2, a representation of $f_{2,1}$ *i.e.* $f_{2,1} = [\langle \mathcal{T}_{2,1} \rangle]_2$.

Given two synchronous sequential transducers $\mathcal{T} = (T, i, \omega)$ and $\mathcal{T}' = (T', i', \omega')$, their *composition* is the following synchronous sequential transducer:

$$\mathcal{T} \circ \mathcal{T}' = (T \circ T', (i, i'), \omega \circ \omega') \text{ where}$$

$$\omega \circ \omega'((p, p')) = v.\omega'(q') \text{ for any } p \in \mathrm{dom}(\omega), \, q' \in \mathrm{dom}(\omega'), \, p' \overset{\omega(p)/v}{\Longrightarrow}_{T'} q'.$$

For instance, the transducer $\mathcal{T}_{2,1}^2$ is represented as follows:

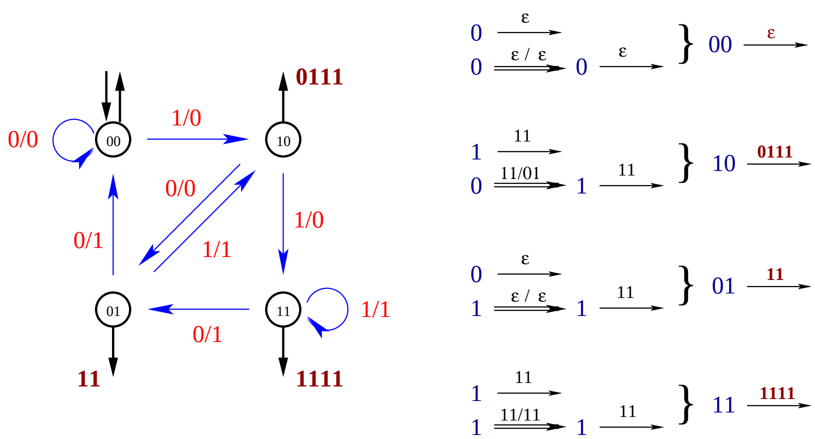


Fig. 15. The composition twice of the previous synchronous sequential transducer.

and realizes the function

$$f_{2,1}^2(n) \;=\; \begin{cases} \frac{n}{4} & \text{if } \; n \in 4\,\mathbb{N} \\ n+1 & \text{if } \; n \in 4\,\mathbb{N} \,+\, 2 \\ 4n+3 & \text{if } \; n \; \text{is odd.} \end{cases}$$

For any $0 \le b < a$ of same parity, we can realize the shortcut $f'_{a,b}$ by adding a terminal function to the transducer of the division by 2 in base a.

Proposition 1. *For all* $0 \le b < a$ *with* $a > 1$ *and* a, b *of same parity,* $\mathcal{T}'_{a,b} \;=\; (:_{a,2}, 0, \omega'_{a,b})$ *with* $\omega'_{a,b}(0) = \varepsilon$ *and* $\omega'_{a,b}(1) = \frac{a+b}{2}$ *is a synchronous sequential transducer for a representation in base* a *of* $f'_{a,b}$ *.*

Proof. The relation $:_{a,2}$ is input-deterministic and input-complete: for all $p \in \{0,1\}$ and $b \in \widehat{a}$, there exists a unique transition starting from p of input b. Thus for all $u \in \widehat{a}^*$, there exists a unique $v \in \widehat{a}^*$ and $j \in \{0,1\}$ such that $0 \overset{u/v}{\Longrightarrow}_{:_{a,2}} j$. By Lemma 3, we have $[u]_a \;=\; 2\,[v]_a + j$.
For $j = 0$, $[u]_a$ is even and $[v]_a \;=\; \frac{[u]_a}{2} \;=\; f'_{a,b}([u]_a)$.
For $j = 1$, $[u]_a$ is odd and since $\frac{a+b}{2} < a$, we have

$$[v\omega'_{a,b}(1)]_a = a[v]_a + \tfrac{a+b}{2} = a\,\tfrac{[u]_a - 1}{2} \,+\, \tfrac{a+b}{2} \,=\, \tfrac{a[u]_a + b}{2} \,=\, f'_{a,b}([u]_a). \;◀$$

Thus $\mathcal{T}'_{3,1} \;=\; (:_{3,2}, 0, \omega'_{3,1})$ with $\omega'_{3,1}(0) = \epsilon$ and $\omega'_{3,1}(1) = 2$ realizes in base 3 the shortcut $f'_{3,1}$ of the Collatz function. A representation of $\mathcal{T}'_{3,1}$ is then obtained from that of $/_{3,2,0}$ given in Figure 8 by adding output 2 to vertex 1.

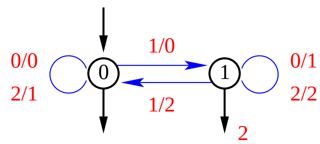


Fig. 16. Shortcut of the Collatz function in base 3.

Translating this transducer in a word rewriting system, we get a variant of the system defined by [13]. Finally, the transducer $\mathcal{T}'_{6,2}$ realizes in base 6 the Collatz function $f_{3,1} = f'_{6,2}$:

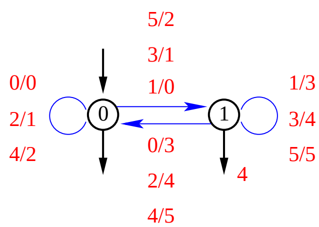


Fig. 17. The Collatz function in base 6.

More generally for any $0 \leq b < a$, the function $f_{a,b} = f'_{2a,2b}$ is represented in base $2a$ by the synchronous sequential transducer $\mathcal{T}'_{2a,2b}$.

5 Transducer for $f'^{p}_{a,b}$ in base a

By composition p times of the previous transducer in base a realizing the shortcut $f'_{a,b}$, we obtain a synchronous sequential transducer realizing $f'^{p}_{a,b}$.

By Lemma 4, the graph of $\mathcal{T}'^{\,p}_{a,b}$ is isomorphic to the division $:_{a,2^p}$ by 2^p in base a: each vertex $x \in \{0,1\}^p$ is in bijection with ${}_2[x] \in \{0,\ldots,2^p-1\}$. So such a transducer realizes the function $f'^{\,p}_{a,b}$, first by doing division by 2^p and then by performing the numerator with a terminal function $\omega_{a,b,p} : \{0,1\}^p \longrightarrow \widehat{a}^*$ to be specified in terms of a, b, p.
For example, below is the 3 times composition of $\mathcal{T}'_{3,1}$ given in Figure 16; it is the transducer of Figure 10 completed with the terminal function $\omega_{3,1,3}$, and which can be compared with that of Figure 5.

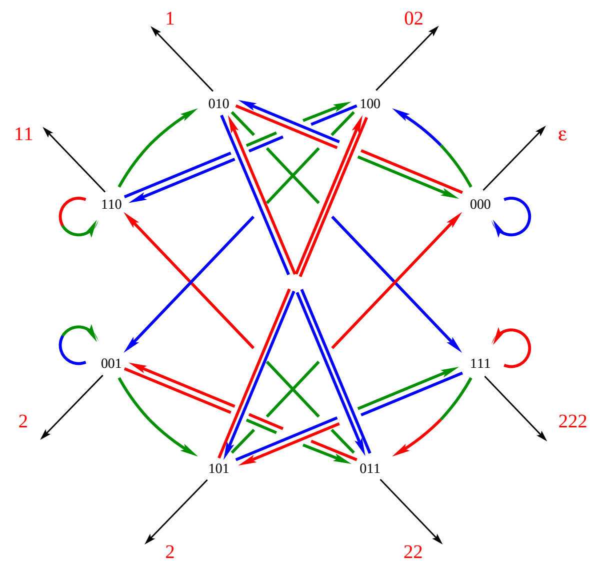


Fig. 18. Another transducer realizing the power of 3 of the Collatz function.

So $\omega_{3,1,3}(100) = 02$. It remains to express $\omega_{3,1,p}(x)$ in terms of p and x.

Lemma 8. *For all $p \geq 0$ and $0 \leq b < a \neq 1$ with a, b of same parity,*

$$[\omega_{a,b,p}(x)]_a = f'^{\,p}_{a,b}({}_2[x])$$
$$a^{|\omega_{a,b,p}(x)|} = f'^{\,p}_{a,b}(2^p + {}_2[x]) - f'^{\,p}_{a,b}({}_2[x]) \quad \textit{for any } x \in \{0,1\}^p.$$

Proof.
i) Let us check the first equality.
Let $u \in \widehat{a}^*$ such that $[u]_a = {}_2[x]$.
As $[u]_a < 2^p$, we have $0^p \overset{u/0^{|u|}}{\Longrightarrow}_{:a,2^p} x \overset{\omega_{a,b,p}(x)}{\longrightarrow}$.
Thus $f'^{\,p}_{a,b}({}_2[x]) = f'^{\,p}_{a,b}([u]_a) = [0^{|u|}\omega_{a,b,p}(x)]_a = [\omega_{a,b,p}(x)]_a$.

ii) Let us check the second equality.
Let $y \in \{0,1\}^p$ and $b \in \widehat{a}$ such that $y \overset{b/1}{\longrightarrow}_{:a,2^p} x$. Thus ${}_2[y]\,a + b = 2^p + {}_2[x]$.
As in (i), let $u \in \widehat{a}^*$ such that $[u]_a = {}_2[y]$.
Hence $0^p \overset{u/0^{|u|}}{\Longrightarrow}_{:a,2^p} y$ and $[ub]_a = a\,[u]_a + b = 2^p + {}_2[x]$. Therefore and by (i),

$$f'^{\,p}_{a,b}(2^p + {}_2[x]) = [1\omega_{a,b,p}(x)]_a = f'^{\,p}_{a,b}({}_2[x]) + a^{|\omega_{a,b,p}(x)|}. \blacktriangleleft$$

To determine $\omega_{a,b,p}(x)$, we compute $[\omega_{a,b,p}(x)]_a$ but also $|\omega_{a,b,p}(x)|$ because possible zeros on the left are significant for a terminal function.

Using the property below [1], we will express this length as $\eta_{a,b,p}({}_2[x])$ where

$$\eta_{a,b,p}(n) \;=\; |\{\; 0 \le i < p \mid f'^{\,i}_{a,b}(n) \text{ odd } \;\}|$$

is the number of odd integers (rises) among the first p powers of $f'_{a,b}$ applied from n. For instance $\eta_{3,1,3}(1) = 2$ since the first three powers of $f'_{3,1}$ starting from 1 are given by the cycle $1 \longrightarrow 2 \longrightarrow 1$.
Note that for any integers q and r,

$$f'_{a,b}(2q+r) \;=\; \begin{cases} q + \frac{r}{2} & \text{if } \; r \text{ is even} \\ a\,q + \frac{ar+b}{2} & \text{otherwise} \end{cases}$$

hence $f'_{a,b}(2q+r) \;=\; q\,a^{\eta_{a,b,1}(r)} + f'_{a,b}(r)$.
This equality has been extended to powers of $f'_{a,b}$ [1].

Lemma 9. *For all natural numbers a, b, p, q, r with a, b of same parity,*

$$f'^{\,p}_{a,b}(q2^p + r) \;=\; q\,a^{\eta_{a,b,p}(r)} + f'^{\,p}_{a,b}(r) \quad \textit{and} \quad \eta_{a,b,p}(q2^p + r) \;=\; \eta_{a,b,p}(r).$$

Proof. By induction on $p \ge 0$.
$p = 0$: $\eta_{a,b,0}$ is the constant mapping 0 and $f'^{\,0}_{a,b}$ is the identity.
$p \Longrightarrow p+1$: For r even, we have

$$\begin{aligned} f'^{\,p+1}_{a,b}(q2^{p+1}+r) \;&=\; f'^{\,p}_{a,b}(f'_{a,b}(q2^{p+1}+r)) \;=\; f'^{\,p}_{a,b}(q2^p + \tfrac{r}{2}) \\ &=\; q\,a^{\eta_{a,b,p}(\frac{r}{2})} + f'^{\,p}_{a,b}(\tfrac{r}{2}) \;=\; q\,a^{\eta_{a,b,p+1}(r)} + f'^{\,p+1}_{a,b}(r) \end{aligned}$$

and $\eta_{a,b,p+1}(q2^{p+1}+r) \;=\; \eta_{a,b,p}(q2^p + \frac{r}{2}) = \eta_{a,b,p}(\frac{r}{2}) = \eta_{a,b,p+1}(r)$.
For r odd, we have

$$\begin{aligned} f'^{\,p+1}_{a,b}(q2^{p+1}+r) &= f'^{\,p}_{a,b}(f'_{a,b}(q2^{p+1}+r)) \;=\; f'^{\,p}_{a,b}(aq2^p + \tfrac{ar+b}{2}) \\ &= f'^{\,p}_{a,b}(aq2^p + f'_{a,b}(r)) \;=\; q\,a^{1+\eta_{a,b,p}(f'_{a,b}(r))} + f'^{\,p}_{a,b}(f'_{a,b}(r)) \\ &= q\,a^{\eta_{a,b,p+1}(r)} + f'^{\,p+1}_{a,b}(r) \end{aligned}$$

and

$$\begin{aligned} \eta_{a,b,p+1}(q2^{p+1}+r) &= 1 + \eta_{a,b,p}(f'_{a,b}(q2^{p+1}+r)) = 1 + \eta_{a,b,p}(aq2^p + f'_{a,b}(r)) \\ &= 1 + \eta_{a,b,p}(f'_{a,b}(r)) \qquad = \eta_{a,b,p+1}(r). \;◀ \end{aligned}$$

From Lemmas 8 and 9, it follows that

$$|\omega_{a,b,p}(x)| \;=\; \eta_{a,b,p}({}_2[x]) \;\text{ for any }\; x \in \{0,1\}^p.$$

Lemma 9 is illustrated below for an accepting path of $\mathcal{T}'^{\,p}_{a,b}$.

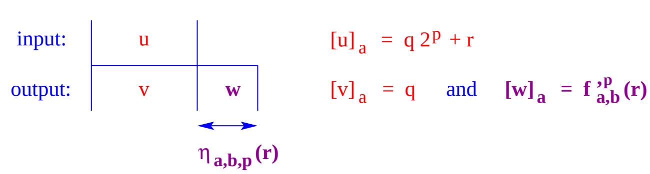


Fig. 19. Terminal function of $\mathcal{T}'^{\,p}_{a,b}$.

We denote by ${}_2[\mathcal{T}'^{\,p}_{a,b}]$ the transducer where each vertex $x \in \{0,1\}^p$ is replaced by the integer ${}_2[x]$. Then ${}_2[\mathcal{T}'^{\,p}_{a,b}]$ is the transducer of division by 2^p in base a with a terminal function defined by the 2^p first values of $f'^{\,p}_{a,b}$. For any vertex, the length of its final word is the number of odd numbers among the first p values of its orbit.

Theorem 1. *For all $p \geq 0$ and $0 \leq b < a \neq 1$ with a, b of same parity, the function $f'^{\,p}_{a,b}$ is recognized by the transducer*

$$ {}_2[\mathcal{T}'^{\,p}_{a,b}] \;=\; (:_{a,2^p}, 0, \omega'_{a,b,p}) $$

of division by 2^p in base a with for any $0 \leq i < d^p$, $\omega_p(i) \in \widehat{a}^*$ is defined by

$$ [\omega'_{a,b,p}(i)]_a \;=\; f'^{\,p}_{a,b}(i) \quad \text{and} \quad |\omega'_{a,b,p}(i)| \;=\; \eta_{a,b,p}(i). $$

Proof. Let us give another proof of this theorem which will be useful later.
By induction on $p \geq 0$. We denote $\omega'_{a,b,p}$ by ω'_p.

$p = 0$: $\mathcal{T}'^{\,0}_{a,b} \;=\; (\{\, \varepsilon \xrightarrow{c/c} \varepsilon \mid c \in \widehat{a}\,\}, \varepsilon, \omega)$ with $\omega(\varepsilon) \;=\; \varepsilon$.

$p \implies p+1$: we have $\mathcal{T}'^{\,p+1}_{a,b} \;=\; \mathcal{T}'_{a,b} \circ \mathcal{T}'^{\,p}_{a,b}$.
By Lemma 4, the relation ${}_2[\,:_{a,2} \circ :_{a,2^p}]$ is equal to $:_{a,2^{p+1}}$.
We have to show that ω'_{p+1} is the terminal function of ${}_2[\mathcal{T}'^{\,p+1}_{a,b}]$.
As $\omega'_{a,b}(0) \;=\; \varepsilon$, we get $\omega'_{p+1}({}_2[0u]) \;=\; \omega'_p({}_2[u])$ for any $u \in \{0,1\}^p$ hence

$$ \omega'_{p+1}(2i) \;=\; \omega'_p(i) \;\text{ for all }\; 0 \leq i < 2^p. \tag{1} $$

By induction hypothesis, we get

$$ [\omega'_{p+1}(2i)]_a \;=\; [\omega'_p(i)]_a \;=\; f'^{\,p}_{a,b}(i) \;=\; f'^{\,p+1}_{a,b}(2i) \quad \text{and} $$
$$ |\omega'_{p+1}(2i)| \;=\; |\omega'_p(i)| \;=\; \eta_{a,b,p}(i) \;=\; \eta_{a,b,p+1}(2i). $$

Similarly $\omega'_{a,b}(1) \;=\; \frac{a+b}{2}$ and for any $0 \leq i < 2^p$, there exists unique j and c such that $i \xrightarrow[:_{a,2^p}]{\frac{a+b}{2}/c} j$ thus

$$ \omega'_{p+1}(2i{+}1) \;=\; c.\omega'_p(j) \;\text{ for all }\; 0 \leq i < 2^p \;\text{ with }\; a\,i + \tfrac{a+b}{2} \;=\; c\,2^p + j. \tag{2} $$

Moreover $f'_{a,b}(2i+1) \;=\; ia + \frac{a+b}{2} \;=\; c2^p + j$.
By Lemma 9 and induction hypothesis,

$$ \begin{aligned} [\omega'_{p+1}(2i+1)]_a \;&=\; [c.\omega'_p(j)]_a \qquad\qquad = c\,a^{|\omega'_p(j)|} + [\omega'_p(j)]_a \\ &=\; c\,a^{\eta_{a,b,p}(j)} + f'^{\,p}_{a,b}(j) \;=\; f'^{\,p}_{a,b}(c2^p + j) \\ &=\; f'^{\,p+1}_{a,b}(2i+1) \end{aligned} $$

and

$$ \begin{aligned} |\omega'_{p+1}(2i+1)| \;&=\; 1 + |\omega'_p(j)| \qquad\qquad = 1 + \eta_{a,b,p}(j) \\ &=\; 1 + \eta_{a,b,p}(c2^p + j) \;=\; 1 + \eta_{a,b,p}(f'_{a,b}(2i+1)) \\ &=\; \eta_{a,b,p+1}(2i+1). \;\blacktriangleleft \end{aligned} $$

Thus the terminal function $\omega'_{a,b,p}(i)$ of any vertex $0 \leq i < 2^p$ is fully determined by the prefix of length p of its orbit:

$$ \underbrace{i \longrightarrow f'_{a,b}(i) \longrightarrow \ldots \longrightarrow f'^{\,p-1}_{a,b}(i)}_{|\omega'_{a,b,p}(i)| \;=\; \text{number of odd integers}} \longrightarrow f'^{\,p}_{a,b}(i) \;=\; [\omega'_{a,b,p}(i)]_a $$

Here is a representation in base 5 of $f'^{\,3}_{5,1}$ by the transducer $\mathcal{T}'^{\,3}_{5,1}$:

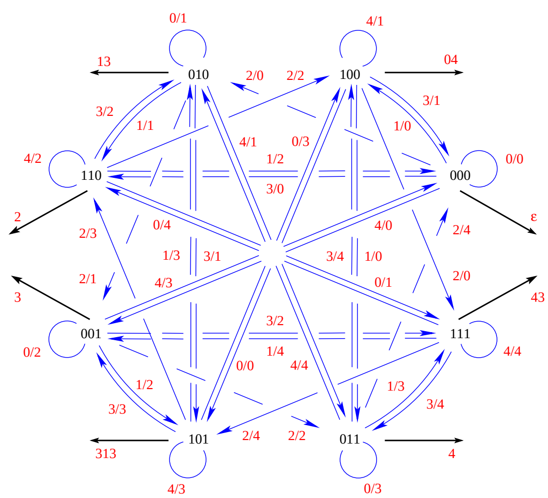


Fig. 20. Transducer realising $f'^{3}_{5,1}$ in base 5.

Finally, the function $f^p_{a,b}$ is realized in base $2a$ by the synchronous sequential transducer $\mathcal{T}^p_{a,b} = \mathcal{T}'^p_{2a,2b}$.

6 Transducer for $f^p_{a,b,d}$ in base ad

For all natural numbers a, b, d with $d \neq 0$, we consider the functions $f_{a,b,d}$: $\mathbb{N} \longrightarrow \mathbb{N}$ defined for any integer $n \geq 0$ by

$$f_{a,b,d}(n) \ = \ \begin{cases} \frac{n}{d} & \text{if} \quad n \text{ is a mutiple of } d, \\ an+b & \text{otherwise.} \end{cases}$$

So $f_{a,b} = f_{a,b,2}$. For $b < a$, we generalize the previous transducers realizing $f_{a,b}$. We define a synchronous sequential transducer realizing $f_{a,b,d}$ from the transducer computing division by d in base ad.

Proposition 2. *For all $0 \leq b < a \neq 1$ and $d > 0$, the synchronous sequential transducer*

$$\mathcal{T}_{a,b,d} \ = \ (:_{ad,d}, 0, \omega_{a,b})$$
$$\textit{with} \quad \omega_{a,b}(0) = \varepsilon \quad \textit{and} \quad \forall\ 0 < j < d,\ \omega_{a,b}(j) \ = \ aj+b$$

realizes a representation in base ad of $f_{a,b,d}$.

Proof.
If an initial path ends to the state j, the input represents an integer n multiple of d plus j and the output represents $\frac{n-j}{d}$. The final digit in $j \neq 0$ is $aj+b$ since $ad\,\frac{n-j}{d} + aj + b \ = \ an+b \ = \ f_{a,b,d}(n)$. ◀

For all integers $a, b, d, p, n \geq 0$ with $d \neq 0$, $\eta_{a,b,p}(n)$ is generalized to the number

$$\mu_{a,b,d,p}(n) \ = \ |\{ \ 0 \leq i < p \mid f^{\,i}_{a,b,d}(n) \ \text{not multiple of } d \ \}|$$

of integers that are not multiples of d among the first p numbers of the orbit from n of $f_{a,b,d}$. Let us adapt Lemma 9 to the powers of $f_{a,b,d}$.

Lemma 10. *For all natural numbers a, b, d, p, q, r with $d \neq 0$, we have* $f^{\,p}_{a,b,d}(qd^p + r) \ = \ q\,(ad)^{\mu_{a,b,d,p}(r)} + f^{\,p}_{a,b,d}(r)$ *and* $\mu_{a,b,d,p}(qd^p + r) \ = \ \mu_{a,b,d,p}(r)$.

Proof. By induction on $p \geq 0$.
$p = 0$: immediate because $\mu_{a,b,d,0}$ is the constant mapping 0 and $f^0_{a,b,d}$ is the identity.
$p \Longrightarrow p+1$: For r multiple of d, we have

$$\begin{aligned} f^{\,p+1}_{a,b,d}(qd^{p+1}+r) &= f^{\,p}_{a,b,d}(f_{a,b,d}(qd^{p+1}+r)) \qquad = f^{\,p}_{a,b,d}(qd^p + \tfrac{r}{d}) \\ &= q\,(ad)^{\mu_{a,b,d,p}(\frac{r}{d})} + f^{\,p}_{a,b,d}(\tfrac{r}{d}) \ = \ q\,(ad)^{\mu_{a,b,d,p+1}(r)} + f^{\,p+1}_{a,b,d}(r) \end{aligned}$$

and

$$\mu_{a,b,d,p+1}(qd^{p+1}+r) = \mu_{a,b,d,p}(qd^p + \tfrac{r}{d}) = \mu_{a,b,d,p}(\tfrac{r}{d}) = \mu_{a,b,d,p+1}(r).$$

For r not multiple of d, we have

$$\begin{aligned} f^{\,p+1}_{a,b,d}(qd^{p+1}+r) &= f^{\,p}_{a,b,d}(f_{a,b,d}(qd^{p+1}+r)) \\ &= f^{\,p}_{a,b,d}(aqd^{p+1}+ar+b) \\ &= f^{\,p}_{a,b,d}((qad)d^p \ + \ f_{a,b,d}(r)) \\ &= qad\,(ad)^{\mu_{a,b,d,p}(f_{a,b,d}(r))} \ + \ f^{\,p}_{a,b,d}(f_{a,b,d}(r)) \\ &= q\,(ad)^{\mu_{a,b,d,p+1}(r)} + f^{\,p+1}_{a,b,d}(r) \end{aligned}$$

and

$$\begin{aligned}\mu_{a,b,d,p+1}(qd^{p+1}+r) &= 1+\mu_{a,b,d,p}(aqd^{p+1}+ar+b)\\ &= 1+\mu_{a,b,d,p}((qad)d^p+f_{a,b,d}(r))\\ &= 1+\mu_{a,b,d,p}(f_{a,b,d}(r))\\ &= \mu_{a,b,d,p+1}(r). \blacktriangleleft\end{aligned}$$

Similarly to Theorem 1, we get an explicit description of the transducer $\mathcal{T}^p_{a,b,d}$ realizing $f^p_{a,b,d}$ for all p.

Theorem 2. *For all integers $p \geq 0$ and $0 \leq b < a \neq 1$ and $d > 0$, the function $f^p_{a,b,d}$ is realized by the synchronous sequential transducer*

$$_d[\mathcal{T}^p_{a,b,d}] \;=\; (:_{ad,d^p}, 0, \omega_{a,b,d,p})$$

with for any $0 \leq i < d^p$, the word $\omega_p(i)$ over $\{0,\ldots,ad-1\}$ is defined by

$$[\omega_{a,b,d,p}(i)]_{ad} \;=\; f^p_{a,b,d}(i) \quad \textit{and} \quad |\omega_{a,b,d,p}(i)| \;=\; \mu_{a,b,d,p}(i).$$

Proof. By induction on $p \geq 0$. We denote $\omega_{a,b,d,p}$ by ω_p.

$p=0$: $\mathcal{T}^0_{a,b,d} \;=\; (\{\,\varepsilon \xrightarrow{c/c} \varepsilon \mid c \in \widehat{ad}\,\}, \varepsilon, \omega)$ with $\omega(\varepsilon) \;=\; \varepsilon$.

$p \implies p+1$: we have $\mathcal{T}^{p+1}_{a,b,d} \;=\; \mathcal{T}_{a,b,d} \circ \mathcal{T}^p_{a,b,d}$.

By Lemma 4, the transition relation $_d[\,:_{ad,d} \circ :_{ad,d^p}]$ is equal to $:_{ad,d^{p+1}}$.

We have to show that ω_{p+1} is the terminal function of $_d[\mathcal{T}^{p+1}_{a,b,d}]$.

As $\omega_{a,b}(0) \;=\; \varepsilon$, we get $\omega_{p+1}(_d[0u]) \;=\; \omega_p(_d[u])$ for any $u \in \widehat{d^p}$ *i.e.*

$$\omega_{p+1}(di) \;=\; \omega_p(i) \;\text{ for all }\; 0 \leq i < d^p.$$

By induction hypothesis, we get

$$\begin{aligned}[\omega_{p+1}(di)]_{ad} \;&=\; [\omega_p(i)]_{ad} \;=\; f^p_{a,b,d}(i) \quad = f^{p+1}_{a,b,d}(di)\\ \text{and}\quad |\omega_{p+1}(di)| \;&=\; |\omega_p(i)| \quad = \mu_{a,b,d,p}(i) \;=\; \mu_{a,b,d,p+1}(di).\end{aligned}$$

Let $0 \leq i < d^p$ and $0 < j < d$. So $\omega_{a,b}(j) \;=\; aj+b \;\leq\; a(d-1)+b \;<\; ad$.

There exists unique k and c such that $i \xrightarrow[:_{ad,d^p}]{aj+b/c} k$ thus $\omega_{p+1}(di+j) \;=\; c.\omega_p(k)$.

Moreover $f_{a,b,d}(di+j) \;=\; iad+aj+b \;=\; cd^p+k$.

By Lemma 10 and induction hypothesis,

$$\begin{aligned}[\omega_{p+1}(di+j)]_{ad} &= [c.\omega_p(k)]_{ad}\\ &= c\,(ad)^{|\omega_p(k)|}+[\omega_p(k)]_{ad}\\ &= c\,(ad)^{\mu_{a,b,d,p}(k)}+f^p_{a,b,d}(k)\\ &= f^p_{a,b,d}(cd^p+k)\\ &= f^{p+1}_{a,b,d}(di+j)\end{aligned}$$

and

$$\begin{aligned}|\omega_{p+1}(di+j)| &= 1+|\omega_p(k)|\\ &= 1+\mu_{a,b,d,p}(k)\\ &= 1+\mu_{a,b,d,p}(cd^p+k)\\ &= 1+\mu_{a,b,d,p}(f_{a,b,d}(di+j))\\ &= \mu_{a,b,d,p+1}(di+j). \blacktriangleleft\end{aligned}$$

7 Transducer for $\mathbf{f^*_{a,b,d}}$

We present a simple infinite synchronous sequential transducer realizing the composition closure $f^*_{a,b,d} = \bigcup_{p\geq 0} f^p_{a,b,d}$ of $f_{a,b,d}$ for $b < a \neq 1$ and $d > 0$.
We start by defining a transducer to realize $f'^*_{a,b}$ with $b < a$. We just have to add to the composition closure $:^*_{a,2}$ of the division by 2 in base a, the set 0^* of initial states and a terminal function defined according to b by length induction that is from the vertex set $\widehat{d}^p$ of $:^p_{a,2}$ into the vertex set $\widehat{d}^{p-1}$ of $:^{p-1}_{a,2}$.

Proposition 3. *For all* $0 \leq b < a$ *with* $a > 1$ *and* a, b *of the same parity, the relation* $f'^*_{a,b}$ *is realized by the transducer*

$$\mathcal{T}'^*_{a,b} \;=\; (:^*_{a,2}\,, 0^*, \omega'_{a,b}) \text{ where for all } u \in \{0,1\}^*,$$

$$\omega'_{a,b}(0u) \;=\; \omega'_{a,b}(u) \quad \text{and} \quad \omega'_{a,b}(1u) \;=\; c.\omega'_{a,b}(v) \quad \text{for} \quad 1u \xrightarrow{b/c}_{:^*_{a,2}} 0v.$$

Proof.
Equations 1 and 2 in the proof of Theorem 1 stipulate that for all $p \geq 0$, the terminal function ω'_{p+1} of $\mathcal{T}'^{p+1}_{a,b}$ is defined recursively for all $0 \leq i < 2^p$ by

$$\omega'_{p+1}(2i) \;=\; \omega'_p(i)$$

$$\omega'_{p+1}(2i+1) \;=\; c.\omega'_p(j) \quad \text{for} \quad i \xrightarrow{\frac{a+b}{2}/c}_{:_{a,2^p}} j$$

and we have

$$\begin{aligned} i \xrightarrow{\frac{a+b}{2}/c}_{:_{a,2^p}} j &\iff ia + \tfrac{a+b}{2} \;=\; c2^p + j \\ &\iff (2i+1)a + b \;=\; c2^{p+1} + 2j \\ &\iff 2i+1 \xrightarrow{b/c}_{:_{a,2^{p+1}}} 2j. \blacktriangleleft \end{aligned}$$

We visualize $\mathcal{T}'^*_{a,b}$ by a cone with ε at the tip and circular sections.

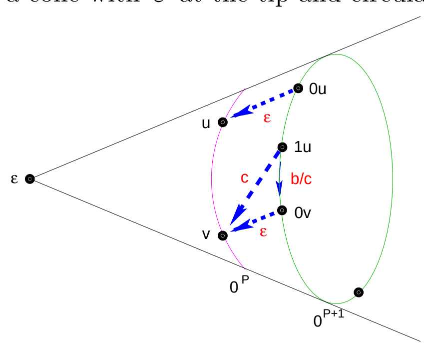


Fig. 21. The composition closure $f'^*_{a,b}$ in base a.

The p-th section is the previously given representation of the Euclidean division $:_{a,2}^{p}$ of initial state 0^p. The terminal function $\omega'_{a,b}$ is represented as follows: with a transition $0u \xrightarrow{\varepsilon} u$ from any node starting by 0, and a transition $1u \xrightarrow{c} v$ from any node starting by 1 for the transition $1u \xrightarrow{b/c} 0v$ of the division by $2^{|u|+1}$ in base a. Note that these transitions of the terminal function can only be used at the end of an accepting path.
Similarly to Proposition 3, we get an explicit description of a transducer realizing $f_{a,b,d}^{*}$ for $b < a$.

Theorem 3. *For all integers $0 \le b < a \ne 1$ and $d > 0$, the relation $f_{a,b,d}^{*}$ is realized by the transducer*

$$\mathcal{T}_{a,b,d}^{*} \;=\; (:_{ad,d}^{*}, 0^{*}, \omega_{a,b,d}) \text{ with for all } u \in \widehat{d}^{\,*} \text{ and } 0 < i < d,$$

$$\omega_{a,b,d}(0u) \;=\; \omega_{a,b,d}(u) \;\text{ and }\; \omega_{a,b,d}(iu) \;=\; c.\omega_{a,b,d}(v) \;\text{ for } iu \xrightarrow{bd/c}_{:_{ad,d}^{*}} 0v.$$

Theorem 3 states that under the condition $b < a \ne 1$, we realize the composition closure of $f_{a,b,d}$ by taking the union $:_{ad,d}^{*}$ of the divisions $:_{ad,d}^{p}$ of initial states 0^p, plus a recurrent terminal function.

8 Conclusion

This work focuses on the description of functions on integers and their powers by deterministic transducers. This has been possible for the functions $f_{a,b,d}$ by the choice of the base ad but only under the restriction that $b < a$. The generalization to any integers a and b requires a new approach.

For any natural numbers a, b, d with $b < a \ne 1$ and $d \ne 0$, we have given an explicit construction of a transducer realizing the closure under composition of $f_{a,b,d}$. In its geometric representation, the disposition of the vertices is well appropriate for both the transitions of the Euclidean divisions and those of the terminal function. It might be a new approach to consider the circularity of the functions $f_{a,b,d}$ namely the existence of paths $0^p \overset{uv/0^{|v|}u}{\Longrightarrow} x$ where v is the terminal word of the vertex x in the transducer of the division by d^p in base ad. However, the circularity of the Collatz function is already considered as a difficult subproblem of the Collatz conjecture.

References

1. J.-P. Allouche, *T. Tao et la conjecture de Syracuse*, Gazette de la SMF 168, 34–39 (2021).
2. J.-P. Allouche, J. Shallit, *Automatic sequences: theory, applications, generalizations*, Cambridge University Press, 588 pages (2003). doi: 10.1017/CBO9780511546563
3. S. Eliahou, J. Fromentin, R. Simonetto, *Is the Syracuse Falling Time Bounded by 12?*, M.B. Nathanson (eds), Combinatorial and Additive Number Theory V. CANT, Springer Proceedings in Mathematics & Statistics, vol. 395, 139–152 (2022). doi: 10.48550/arXiv.2107.11160
4. P. Erdös, *Some Unconventional Problems in Number Theory*, Mathematics Magazine, vol. 52 (2), 67–70 (1979). doi : 10.2307/2689842
5. J. Lagarias, *The ultimate challenge: the 3x+1 problem*, American Mathematical Society (2010).doi: 10.1090/mbk/078
6. J.E. Pin, *Variétés de langages formels*, Masson, Paris (1984), and *Varieties of formal languages*, North Oxford, London and Plenum, New-York (1986). doi: 10.1137/1031081
7. G.W. Reitwiesner, *Binary Arithmetic*, Advances in Computers, vol. 1, New York Academic Press , 231-308 (1960). doi: 10.1016/S0065-2458(08)60610-5
8. M.-P. Schützenberger, *Sur une variante des fonctions séquentielles*, Theoretical Computer Science 4 (1), 47–57 (1977) doi: 10.1016/0304-3975(77)90055-X
9. J. Shallit and D. Wilson, *The "$3x + 1$" Problem and Finite Automata*, Bulletin of the EATCS, No. 46, 182–185 (1992).
10. P. Simonnet, *Personal communication* (2019).
11. T. Stérin, D. Woods, *Binary expression of ancestors in the Collatz graph*, S. Schmitz and I. Potapov (eds), Reachability Problems, Lecture Notes in Computer Science, vol. 12448, Springer, 131-147 (2020).
12. T. Tao, *Almost all orbits of the Collatz map attain almost bounded values*, arXiv 1909.03562 (2019).doi: 10.48550/arXiv.1909.03562
13. E. Yolcu, S. Aaronson, M.J.H. Heule, *An Automated Approach to the Collatz Conjecture*, A. Platzer and G. Sutcliffe (eds), Automated Deduction, 28th CADE, Lecture Notes in Computer Science, vol. 12699, Springer, 468-484 (2021). doi: 10.48550/arXiv.2105.14697