Linear numeration systems, $\theta$-developments and finite automata

Christiane Frougny
LITP and Université René Descartes
UER de Mathématiques et Informatique
12, Rue Cujas, 75005 Paris, France

ABSTRACT

In numeration systems defined by a linear recurrence relation, as well as in the set of developments of numbers in a non integer basis $\theta$, we define the notion of normal representation of a number. We show that, taking for $\theta$ the greatest root of the characteristic polynomial of the linear recurrence, and under certain conditions of confluence, the normal representation can be obtained from any representation by a finite automaton which is the composition of two sequential transducers derived from the linear recurrence. The addition of two numbers can be performed by a left sequential transducer.

1. Introduction

The representation of a number is the writing of this number as a word on an alphabet - the alphabet of the digits - with respect to some basis. In this paper we study numeration systems where the basis is defined by a linear recurrence. The representation of an integer obtained by the usual algorithm is called the normal representation of that integer. It is a word written on the canonical alphabet. We show that if the relation defined by the linear recurrence is confluent, this normal representation can be computed from any representation by a finite automaton which is the composition of two subsequential transducers linked to the linear recurrence. We call this process the normalization. The addition of two numbers can be viewed as a particular case of a normalization generalized to an arbitrary set of digits. The transformation of a word written on an arbitrary set of digits into a word on the canonical alphabet having the same numerical value can be performed by a subsequential transducer.

The real numbers are usually represented by their $\theta$-development where $\theta > 1$ is a real number non integer ([11]). The $\theta$-shift is the set of infinite sequences which are $\theta$-developments of the elements of $[0,1]$. The properties of the $\theta$-shift have been largely studied (cf. [10] and [5], [4] where references on these topics can be found).

This research has been partly supported by the Programme de Recherches Coordonnées Mathématiques et Informatique.
In this paper, we consider not only the $\theta$-development but any $\theta$-representation of a real number in basis $\theta$. A $\theta$-development is the normal form of a $\theta$-representation of a real number. Only the case where the $\theta$-development of 1 is finite is examined. The relationship between linear numeration systems and $\theta$-shift has been explicated in [4]. The representation of an integer in a linear system is a finite word on an alphabet, and the $\theta$-representation of a real number is an infinite word. Hence to study the normalization process for $\theta$-representations, we introduce the notion of the infinite behavior of a transducer. The transducers considered in the infinite case have the same underlying finite transducer than in the finite case. We prove that the $\theta$-development of a real number can be computed from any $\theta$-representation by the composition of two sequential transducers, and that the addition is performed by a sequential transducer.

In order to state the results more precisely let us give some definitions.

First consider the representation of positive integers. Given a strictly increasing sequence $U = (u_n)_{n \geq 0}$ of non-negative integers, with $u_0 = 1$, every positive integer $N$ can be written in this basis $U$, that is one can find $n \geq 0$ and integers $d_0, \ldots, d_n$ such that $N = d_0 u_n + \cdots + d_n u_0$. $d_0 \cdots d_n$ is called a representation of $N$ in the basis $U$.

That can be done by the following algorithm (folklore). Given integers $x$ and $y$, denote by $q(x,y)$ and $r(x,y)$ the quotient and the rest of the euclidean division of $x$ by $y$. Let $n \geq 0$ such that $u_n \leq N < u_{n+1}$ and let $d_0 = q(N,u_n)$ and $r_0 = r(N,u_n)$. Iterate this process, that is compute $d_i = q(r_{i-1},u_{n-i})$, $r_i = r(r_{i-1},u_{n-i})$ for $i = 1,\ldots,n$. Then, for $0 \leq i \leq n$, $d_i < \frac{u_{n-i+1}}{u_{n-i}}$. Thus, if there exists a positive constant $K$ such that for every $n$, $\frac{u_{n+1}}{u_n} \leq K$, $K$ minimum, then $0 \leq d_i \leq K-1$. We call $A = \{0,\ldots,K-1\}$ the canonical alphabet of digits associated to $U$. The (unique) representation $d_0 \cdots d_n$ obtained by this algorithm is called the normal representation of $N$.

Fact 1. The normal representation of an integer (with respect to some basis $U$) is greater, in the lexicographic ordering, than any other representation of the same length of that integer.

In this paper, the basis $U$ is defined by a linear recurrence relation of order $m \geq 2$:

$$u_{n+m} = a_1 u_{n+m-1} + \cdots + a_m u_n$$

$a_i \in \mathbb{N}$, $u_0 = 1$, $u_1, \ldots, u_{m-1}$ given

such that $U$ is strictly increasing. The system $U$ is called a linear numeration system.

Example 1. The Fibonacci numeration system is defined by

$$u_{n+2} = u_{n+1} + u_n$$

$u_0 = 1$, $u_1 = 2$