ELSEVIER Linear Algebra and its Applications 351–352 (2002) 91–98

LINEAR ALGEBRA AND ITS APPLICATIONS

www.elsevier.com/locate/laa

The presence of a zero in an integer linear recurrent sequence is NP-hard to decide

Vincent D. Blondel [a,*], Natacha Portier [b]

[a] *Division of Applied Mathematics, Center CESAME, Université catholique de Louvain, 4 Avenue Georges Lemaitre, B-1348 Louvain-la-Neuve, Belgium*

[b] *Laboratoire de l'Informatique du Parallélisme, École normale Supérieure de Lyon, 46 Allée d'Italie, F-69364 Lyon Cédex 07, France*

Received 15 August 2000; accepted 31 July 2001

Submitted by D. Hinrichsen

Abstract

We show that the problem of determining if a given integer linear recurrent sequence has a zero—a problem that is known as "Pisot's problem"—is NP-hard. With a similar argument we show that the problem of finding the minimal realization dimension of a one-letter max-plus rational series is NP-hard. This last result answers a folklore question raised in the control literature on the max-plus approach to discrete event systems. Our results are simple consequences of a construction due to Stockmeyer and Meyer.

AMS classification: 93B20; 93C65; 05C50

Keywords: Pisot's problem; Linear recurrent sequence; Minimal realization; Max-plus rational series

0. Introduction

We show that the problem of determining if a given integer linear recurrent sequence has a zero is NP-hard. It is not known if problem is decidable. With a similar argument we show that the problem of finding the minimal realization dimension of a one-letter max-plus rational series is NP-hard. This last result answers a folklore question raised in the control literature on the max-plus approach to discrete event

* Corresponding author. Tel.: +32-10-472-381; fax: +32-10-472-180.
E-mail addresses: blondel@inma.ucl.ac.be (V.D. Blondel), natacha.portier@ens-lyon.fr (N. Portier).

0024-3795/02/$ - see front matter
PII: S0024-3795(01)00466-9

systems. We study the decidability of this question with Stéphane Gaubert in a later article [3].

These results are simple consequences of a little-known construction given by Stockmeyer and Meyer that shows how to polynomialy reduce the 3SAT satisfiability problem to the problem of determining if a given rational expression [1] over the unary alphabet $\{a\}$ is equal to $(a)^*$.

In this paper, we reproduce the construction of Stockmeyer and Meyer and derive consequences of this construction for graphs, rational series, Pisot's problem, and for the minimal realization problem of max-plus discrete event systems.

1. Rational expressions

The problem of determining if two rational expressions over a unary alphabet define different languages is proved NP-hard in Theorem 6.1 of [27]. The argument used by the authors in fact establishes the stronger statement that the problem remains NP-hard even if one of the rational expressions is equal to $(a)^*$.

Theorem 1.1. *The problem of determining if the language defined by a rational expression over a unary alphabet $\{a\}$ is different from $(a)^*$ is NP-hard.*

Proof. The proof is by reduction from the satisfiability problem 3SAT [10]. Let $S = C_1 \wedge C_2 \wedge \cdots \wedge C_m$ be a conjunction of clauses $C_1, \ldots, C_m$ in the variables $x_1, \ldots, x_n$. Each clause C_k is the disjunction of exactly three terms (a term is a variable x_i or the negation of a variable $\neg x_i$). We construct a rational expression E defining a language $L \subseteq (a)^*$ such that $a^z \in L$ whenever $z \geqslant 0$ is not the code for a solution of S (the coding function is described next). The construction of E is given by a polynomial time algorithm and so it will establish the result.

Let $p_1, \ldots, p_n$ be the first n primes. According to the Prime Number theorem, the number $\pi(i)$ of primes less than or equal to i is of the order $i/\ln i$. Hence we have $p_i \leqslant i^2$ when i is large enough and the first n primes can be computed in time polynomial in n.

To the nonnegative integer z we associate the tuple

$$\mu(z) = \big(z \bmod p_1,\ z \bmod p_2, \ldots, z \bmod p_n\big) \in \mathbf{N}^n,$$

where $(z \bmod p_i)$ denotes the remainder of z modulo p_i. The integer z is said to be a *code* for $\bar{a} = (a_1, a_2, \ldots, a_n) \in \mathbf{N}^n$ if $\mu(z) = \bar{a}$. According to the Chinese Remainder theorem (see, for example, [21, p. 94]), every Boolean tuple $\bar{a} \in \{0, 1\}^n$ has a code.

[1] A rational expression over a unary alphabet $\{a\}$ is an expression involving the empty word 1, the letter a, the star operation $*$ and the concatenation and the union operations $\cdot$ and $\cup$. For example, $1 \cup a((1 \cup a \cup aaa)^* aa)^*$ is a rational expression.

The set E is the union of the sets E_0 and $(E_k)_{1\leqslant k\leqslant m}$ where these sets are defined as follows:

- E_0 is the set of words a^z for which $z \geqslant 0$ is not a code for a Boolean tuple, i.e., $\mu(z) \notin \{0,1\}^n$. An integer z is not a code for a Boolean tuple if z modulo p_k is different from 0 and 1 for some k between 1 and n. This set can be seen equal to
$$E_0 = \bigcup_{1\leqslant k\leqslant n} \bigcup_{2\leqslant j\leqslant p_k-1} a^j (a^{p_k})^*.$$
- E_k is the set of words a^z for which $z \geqslant 0$ is such that the vector $\mu(z)$ has Boolean entries at the indices associated to variables appearing in the clause C_k and the corresponding variable assignment is not a solution for C_k. Suppose $C_k = \alpha_{k_1} x_{k_1} \vee \alpha_{k_2} x_{k_2} \vee \alpha_{k_3} x_{k_3}$, with k_1, k_2 and k_3 integers between 1 and n, and with α_{k_1}, α_{k_2} and α_{k_3} Booleans with the convention that $0x$ means $\neg x$ and $1x$ means x. Then a tuple of Booleans $\bar{a} = (a_1, \ldots, a_n)$ is a solution for C_k if and only if $\alpha_{k_1} = a_{k_1}$, $\alpha_{k_2} = a_{k_2}$ or $\alpha_{k_3} = a_{k_3}$. If z modulo p_{k_1}, p_{k_2} and p_{k_3} are all equal to 0 or 1, then the corresponding variable assignments is not a solution of C_k if and only if $z \bmod p_{k_1} = 1 - \alpha_{k_1}$, $z \bmod p_{k_2} = 1 - \alpha_{k_2}$ and $z \bmod p_{k_3} = 1 - \alpha_{k_3}$. Let z be a particular solution to these three equations. Then the set of all solutions is given by $\{z + i p_{k_1} p_{k_2} p_{k_3} : i \in \mathbf{Z}\}$. There exists a unique integer z_k between 0 and $p_{k_1} p_{k_2} p_{k_3} - 1$ in this set and this integer can be found in polynomial time. The set E_k is then seen equal to
$$E_k = a^{z_k} \left(a^{p_{k_1} p_{k_2} p_{k_3}}\right)^*.$$
Finally, let
$$E = E_0 \cup \bigcup_{1\leqslant k\leqslant m} E_k.$$
The rational expression E defines the language $(a)^*$ if and only if S is not satisfiable. Hence the result. □

According to Kleene's theorem [18,26], a language defined by a rational expression can be recognized by a nondeterministic finite automaton (NFA). Since a rational expression can be transformed in polynomial time into an NFA that recognizes the same language, we deduce the following corollary of Theorem 1.1.

Corollary 1.1. *The problem of determining for a given nondeterministic finite automata over a unary alphabet $\{a\}$ if the language recognized by the automaton is different from $(a)^*$ is NP-hard.*

Let $L \subset \{a\}^*$ be the language recognized by an NFA $\mathscr{A}$ with n states. It is easy to see how to construct an NFA $\mathscr{A}'$ with $n+2$ states and with exactly one input state and one output state that recognizes the language $L\backslash\{1\}$. Such an NFA can also be seen as a finite directed graph with two distinguished vertices V_1, V_2; one for the input state and the other for the output state. The word a^k $(k \geqslant 1)$ is accepted

by the automaton $\mathscr{A}$ if there is a directed path of length k from V_1 to V_2. This correspondence between NFAs over unary alphabets and directed graphs gives us another corollary of Theorem 1.1.

Corollary 1.2. *The problem of determining for a given directed graph G and vertices V_1 and V_2 of G if there are directed paths from V_1 to V_2 of all possible lengths $k \geqslant 1$ is co-NP-hard.*

This last corollary also has a rational series interpretation. Let G be a directed graph and let $\{1, \ldots, n\}$ be the set of vertices of G. The adjacency matrix of G is defined by $A = (a_{i,j}) \in \{0,1\}^{n\times n}$, where $a_{i,j}$ is equal to 1 if there is an edge from i to j in G and is equal to 0 otherwise. Let $b \in \{0,1\}^n$ be the row vector whose components are all equal to 0 except for the first component that is equal to 1. Let $c \in \{0,1\}^n$ be the row vector whose components are all equal to 0, except for the last component that is equal to 1. Then, there is a directed path of length 1 from the vertex 1 to the vertex n in G if and only if $b^{\mathrm{T}}Ac$ is equal to 1. In general the entry on row i and column j of A^k $(k \geqslant 1)$ is positive if and only if there is a directed path of length k from the vertex i to the vertex j. Since the scalars $b^{\mathrm{T}}A^k c$ are nonnegative integers for all $k \geqslant 0$ we obtain the following corollary.

Corollary 1.3. *The problem of determining for a given matrix $A \in \{0,1\}^{n\times n}$ and row vectors $b, c \in \{0,1\}^n$ if $b^{\mathrm{T}}A^k c = 0$ for some $k \geqslant 0$ is NP-hard.*

2. Zeros in linear recurrent sequences

Theorem 1.1 has an immediate consequence for the long-standing problem of determining when an integer linear recurrent sequence has a zero coefficient; a problem that is known as Pisot's problem, see e.g. [2,5,20,24,25]. In 1935, Mahler showed that the set of indices of zero coefficients in a recurrent sequence is the union of a finite set and of a finite number of arithmetic progressions. For linear recurrent sequences of order 3, these sets of indices can be constructed effectively (see [29]) and so Pisot's problem is decidable for sequences of order 3. However, no such effective construction is known for sequences of arbitrary order and it is unknown whether Pisot's problem is decidable or not. We show that the problem is NP-hard. For this purpose, consider the matrices $A \in \{0,1\}^{n\times n}$, $b, c \in \{0,1\}^n$, let $p(s)$ be the characteristic polynomial of A

$$p(s) = \det(sI - A) = s^n - a_{n-1}s^{n-1} - \cdots - a_1 s - a_0$$

and define $\gamma_i = b^{\mathrm{T}}A^i c \geqslant 0$ for $i \geqslant 0$. We have

$$\begin{aligned}
&\gamma_{k+n} - a_{n-1}\gamma_{k+n-1} - \cdots - a_1\gamma_{k+1} - a_0\gamma_k \\
&\quad = b^{\mathrm{T}}A^{k+n}c - a_{n-1}b^{\mathrm{T}}A^{k+n-1}c - \cdots - a_1 b^{\mathrm{T}}A^{k+1}c - a_0 b^{\mathrm{T}}A^k c \\
&\quad = b^{\mathrm{T}}A^k(A^n - a_{n-1}A^{n-1} - \cdots - a_1 A^1 - a_0 A^0)c \\
&\quad = 0.
\end{aligned}$$

For the last equality we have used the fact that a matrix solves its characteristic polynomial. Thus, the sequence $(\gamma_i)_{i\geqslant 0}$ satisfies a linear recurrence equation of order n. Since all transformations given above can be performed in polynomial time we conclude from Corollary 1.3:

Corollary 2.1. *It is NP-hard to decide if a given integer linear recurrent sequence has a zero.*

3. The minimal realization problem

We now describe an implication of Theorem 1.1 for the minimal realization problem for discrete event systems. Discrete event systems are dynamical systems whose dynamics are event-driven. Such systems appear in a wide range of practical situations and there are a large number of papers on the computational complexity of discrete event systems that address issues such as minimal realization [7], reachability [15] and stability [4,28]; see also [30] for a survey.

In general, models that describe the behavior of discrete event systems are nonlinear but there is a class of discrete event systems for which the model becomes linear when formulated in the max-plus semiring, see [1] for more details. In the context of these systems, the minimal realization problem is equivalent to that of finding the minimal realization dimension of a one-letter max-plus rational series.

It is shown in [8] that, in the case of Boolean matrices, a lower bound on the minimal realization dimension is given by a quantity (the max-plus Schein rank) that is NP-hard to compute in general. This result does not have computational complexity implications for the minimal realization dimension of arbitrary max-plus systems and the computation of this minimal realization dimension is listed in [23] as one of the open problems in the field of mathematical control theory. The problem has given rise to a number of interesting contributions in the last decade, see, e.g., [6,9,11,12,14]. We give here a proof that the problem is NP-hard.

For this purpose, let G be a directed graph and let $\{1, \ldots, n\}$ be the set of vertices of G. Consider the Boolean semiring $\mathscr{B} = \{0, 1\}$ for which addition and multiplication are defined as usual except that $1 + 1 = 1$. We construct the adjacency matrix $A = (a_{i,j}) \in \{0, 1\}^{n\times n}$ of the graph in $\mathscr{B}$ by setting the entry $a_{i,j}$ equal to 1 if there is an edge from i to j and equal to 0 otherwise. Then, for all integers $k \geqslant 0$, the entry on line i and column j of A^k is equal to 1 if there is a path of length k from i to j in G, and is equal to 0 otherwise. Using a definition of the vectors $b, c \in \{0, 1\}^n$ identical to the one given in Section 3, Corollary 1.2 can be rephrased as follows:

Corollary 3.1. *The problem of determining for a given matrix $A \in \{0,1\}^{n\times n}$ and vectors $b, c \in \{0,1\}^n$ in the Boolean semiring $\mathscr{B}$ if $b^{\mathrm{T}}A^k c = 0$ for some $k \geqslant 0$ is NP-hard.*

A statement similar to Corollary 3.1 can be obtained for any semiring $\mathscr{K}$ in which the Boolean semiring $\mathscr{B}$ can be embedded. Let $\mathscr{K}$ be a semiring and let $\nu : \mathscr{B} \to \mathscr{K}$ be a map such that $\nu(0) \neq \nu(1)$, $\nu(a+b) = \nu(a) + \nu(b)$ and $\nu(a \times b) = \nu(a) \times \nu(b)$. Define $\mathbf{0}, \mathbf{1} \in \mathscr{K}$ by $\nu(0) = \mathbf{0}$ and $\nu(1) = \mathbf{1}$. By translating the statement of Theorem 3.1 in the context of the semiring $\mathscr{K}$ we see that the problem of determining, for a given matrix $A \in \{\mathbf{0}, \mathbf{1}\}^{n\times n}$ and vectors $b, c \in \{\mathbf{0}, \mathbf{1}\}^n$, if $b^{\mathrm{T}}A^k c = \mathbf{0}$ for some $k \geqslant 0$, is NP-hard. This in particular is true for the tropical min-plus and max-plus semirings. The max-plus semiring is the set $\mathbf{N} \cup \{-\infty\}$ equipped with the operations max and $+$ and the min-plus semiring is the set $\mathscr{M} = \mathbf{N} \cup \{+\infty\}$ equipped with min and $+$. These semirings have numerous applications, in particular in dynamic programming, discrete event system theory, optimal control, and asymptotic analysis, see, e.g., [1,13,16,19,22].

In [17], Kanta and Krob introduce a normalized form for rational series that allows computations with such series; see also [12]. Their approach can be used to prove the decidability of the problem of determining if $b^{\mathrm{T}}A^k c = 0$ for some $k \geqslant 0$ in the semiring $\mathscr{M}$. The Boolean semiring $\mathscr{B}$ can be embedded in $\mathscr{M}$ by defining $\nu(1) = 0$ and $\nu(0) = +\infty$ and thus our result shows that, although decidable, this problem is NP-hard.

The minimal realization dimension of a rational series is equal to 1 if and only if the coefficients of the series are all equal. The following result is an immediate consequence of Corollary 3.1.

Corollary 3.2. *The problem of determining for a given matrix $A \in \{-\infty, 0\}^{n\times n}$ and vectors $b, c \in \{-\infty, 0\}^n$ over the max-plus semiring if there exists a realization of dimension* 1 *for the series with the linear representation (b, A, c) is NP-hard.*

Since it is NP-hard to decide if a realization of dimension one exists, it is also NP-hard to decide if there is a realization of size N when N is a part of the data.

References

[1] F. Baccelli, G. Cohen, G.J. Olsder, J.P. Quadrat, Synchronization and Linearity, Wiley, New York, 1992.
[2] J. Berstel, C. Reutenauer, Rational Series and their Languages, Springer, Berlin, 1988.
[3] V.D. Blondel, S. Gaubert, N. Portier, The set of realizations of a max-plus linear sequence is semi-polyhedral, Preprint.
[4] V.D. Blondel, J.N. Tsitsiklis, Complexity of stability and controllability of elementary hybrid systems, Automatica 35 (1988) 479–489.

[5] L. Cerlienco, M. Mignotte, F. Piras, Suites récurrentes linéaires: propriétés algébriques et arithmétiques, Enseign. Math. 33 (1–2) (1987) 67–108.
[6] R.A. Cuninghame-Green, P. Butkovič, Discrete event systems: the strictly convex case, Ann. Oper. Res. 57 (1995).
[7] B. DasGupta, E.D. Sontag, A polynomial-time algorithm for an equivalence problem which arises in hybrid systems theory, in: Proceedings of the 37th IEEE Conference on Decision and Control, 1998, pp. 1629–1634.
[8] B. De Schutter, V.D. Blondel, R. de Vries, B. De Moor, On the boolean minimal realization problem in the max-plus algebra, Systems Control Lett. 35 (2) (1998) 69–78.
[9] B. De Schutter, B. De Moor, Minimal realization in the max algebra is an extended linear complementary problem, Systems Control Lett. 25 (2) (1995) 103–111.
[10] M.R. Garey, D.S. Johnson, Computers and Intractability: A Guide to the Theory of NP-completeness, Freeman, New York, 1979.
[11] S. Gaubert, Théorie des systèmes linéaires dans les dioides, Thèse de Doctorat, Ecole des Mines de Paris, 1992.
[12] S. Gaubert, On rational series in one variable over certain dioids, Technical Report 2162, INRIA, 1994.
[13] S. Gaubert, M. Plus, Methods and applications of $(\max, +)$ linear algebra, in: R. Reischuk, M. Morvan (Eds.), STACS'97, Lecture Notes in Computer Science, vol. 1200, Lübeck, Springer, Berlin, 1997.
[14] S. Gaubert, R.A. Cuninghame-Green, P. Butkovič, Minimal $(\max, +)$ realization of convex sequences, SIAM J. Control Optim. 36 (1) (1998) 137–147.
[15] C. Golaszewski, P.J. Ramadge, The complexity of some reachability problems for a system on a finite group, Systems Control Lett. 12 (1989) 431–435.
[16] J. Gunawardena (Ed.), Idempotency, Publications of the Newton Institute, Cambridge University Press, Cambridge, 1998.
[17] M. Kanta, D. Krob, One-letter rational series with multiplicities in the tropical semiring: an algorithmic approach, Prépublication du LIAFA, Université Paris 7, 2 place Jussieu, 75 251 Paris Cédex 05, France, 1997.
[18] S.C. Kleene, Representation of events in nerve nets and finite automata, in: Automata Studies, Annals of Mathematical Studies, vol. 34, Princeton University Press, Princeton, 1956, pp. 3–41.
[19] V. Kolokoltsov, V. Maslov, Idempotent Analysis and Applications, Kluwer Academic Publisher, Dordrecht, MA, 1997.
[20] W. Kuich, A. Salomaa, Semirings, Automata, Languages, EATCS Monographs on Theoretical Computer Science, vol. 5, Springer, Berlin, 1986.
[21] S. Lang, Algebra, third edition, Addison-Wesley, Reading, MA, 1997.
[22] V. Maslov, S. Samborski (Eds.), Idempotent Analysis, Advances in Soviet Mathematics, vol. 13, AMS, Providence, RI, 1992.
[23] G.-J. Olsder, B. De Schutter, The minimal realization problem in the max-plus algebra, in: V. Blondel, E. Sontag, M. Vidyasagar, J. Willems (Eds.), Open Problems in Mathematical Systems and Control Theory, Springer, London, 1999.
[24] K. Ruohonen, Zeros of Z-rational functions and DOL-equivalence, Theoret. Comput. Sci. 3 (1976) 282–292.
[25] A. Salomaa, M. Soittola, Automata-Theoretic Aspects of Formal Power Series, Springer, New York, 1978.
[26] M.P. Schützenberger, On the definition of a family of automata, Inform. Control 4 (1961) 245–270.
[27] L.J. Stockmeyer, A.R. Meyer, Word problems requiring exponential time: preliminary report, Fifth Annual ACM Symposium on Theory of Computing, ACM, New York, 1973.
[28] O. Toker, On the algorithmic unsolvability of some stability problems for discrete event systems, Proc. IFAC World Congress (1997) 353–358.

[29] N.K. Vereshchagin, On the zeros of linear recurrence sequences, Soviet Math. Dokl. 30 (2) (1984) 502–505.

Reference added in proof

[30] V.D. Blondel, J.N. Tsitsiklis, A survey of computational complexity results in systems and control, Automatica 36 (9) (2000) 1249–1274.