Matrix representations of trace monoids

Christian Choffrut
LIAFA, Université Paris 7, 2, Pl. Jussieu
75 251 Paris Cedex 05, France
cc@liafa.jussieu.fr

Abstract

We consider trace monoids i.e., free monoids where some pairs of letters are allowed to commute. We show that such monoids can be faithfully represented by 2×2-matrices with integer entries if and only if it they are direct products of a free commutative monoid with a free product of free commutative monoids.

1 Introduction

Representation theory is a major field in finite group theory [6]. It allows to view an abstractly defined group as a group of matrices. The depth of this theory is usually illustrated by observing that some results like the famous theorem of Frobenius on transitive groups have to date, no proof which would not involve the use of characters. Concerning infinite groups, designing homomorphic images of groups helps finding subgroups of groups which are abstractly defined by generators and relators (see e.g., [11]).

There is no such deep results for free monoids yet, however it is well known that they can be faithfully represented by 2×2-matrices with entries in the integers $\mathbb{N}$. This has a clear theoretical interest and has also been used in different contexts to solve apparently completely disconnected problems. Some of them concern a few decidability results on $\mathbb{N}$-rational series (cf. e.g. [7, Thm 12.1]) where Post Correspondence Problem can be reduced to some properties of the coefficients of the series. An even more striking application deals with a conjecture of Ehrenfeucht stating that any system of equations with finitely many unknowns in the free monoid is equivalent to a finite subsystem (cf. [1], also [4]). A major step of the proof of this conjecture relies on the existence of a faithful representation of free monoids by matrices with entries in $\mathbb{N}$.

Here we study a classical extension of the free monoids by allowing certain pairs of generators to commute (the *trace monoids*) and we inquire about the conditions under which these monoids can be faithfully represented by 2×2-matrices. Our main result states the following:

If a trace monoid has a faithful representation by 2×2-matrices with entries in the field of complex numbers then it is the direct product of a free commutative monoid with the free product of free commutative monoids.

Conversely, if it is the direct product of a free commutative monoid with the free product of free commutative monoids then it can be faithfully represented by 2×2-matrices with entries in the natural numbers $\mathbb{N}$.

We now turn to a brief description of the various sections of our work. Section 2 deals with the basic definitions on trace monoids and the standard operations of direct and free products. It recalls the classical representations of free monoids which can be found in the literature, their immediate consequences and the decision issues. Some elementary results on commutativity in the monoid of 2×2-matrices with entries in the integers $\mathbb{N}$ are stated which if used in the "only if" part of the Theorem. The last section is devoted to the other direction of the proof which is essentially obtained by exhibiting a representation of every free product of free commutative monoids by 2×2-matrices.

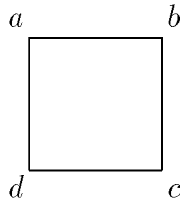
Some part of the present material has already been published in [3]. Some other is new, like the fact that it is recursively decidable whether or not two given 2×2-matrices with entries in $\mathbb{N}$ generate a free commutative monoid.

2 Preliminaries

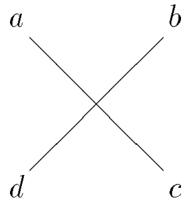
2.1 Free monoids – Partial commutations

The reader is referred to the handbook [9] for a thorough exposition of the theory of traces.

Given a finite nonempty set A called *alphabet* whose elements are *letters*, we denote by A^* the free monoid it generates. An *independence* relation is a non reflexive symmetrical relation $I \subseteq A \times A$ and the corresponding *dependence* relation $\bar{I}$ is the relation consisting of all pairs of distinct elements in A not belonging to I. The congruence on A^* generated by the relators: $ab \sim ba$ for all $(a, b) \in I$ is denoted by $\sim_I$ or simply by $\sim$ when I is understood. The quotient monoid $M(A, I) \equiv A^*/\sim$ is the *trace monoid* generated by A with independence relation I. The free monoid A^* and the free commutative monoid are two extreme cases corresponding to I and $\bar{I}$ respectively being empty. A trace monoid is entirely defined by the graph of its independence relation or equivalently by the graph of its dependence relation. For example if $A = \{a, b, c, d\}$ and if I is the relation $ab \sim ba$, $bc \sim cb$, $cd \sim dc$, $da \sim ad$ then the graphs of I and of $\bar{I}$ are as follows:



The independence relation



The dependence relation

There exist standard ways of defining new partial commutations from a given set of partial commutations which can be interpreted as operations on trace monoids. In particular the free product and the direct product of trace monoids can be defined in terms of simple operations on the independence and the dependence relations. Assume A and B are two disjoint alphabets with dependence relations I and J respectively. Then the trace monoid generated by $A \cup B$ with independence relation $I \cup J$ is the *free product* of the two trace monoids $M(A, I)$ and $M(B, J)$: $M(A \cup B, I \cup J) \equiv M(A, I) * M(B, J)$. Furthermore, the trace monoid generated by $A \cup B$ with the dependence relation $\bar{I} \cup \bar{J}$ is the direct product of the two trace monoids $M(A, I)$ and $M(B, J)$. E.g., the above example is the direct product of two free monoids with two generators.

The following notion is crucial to our purpose. Given a monoid M, we say it is *k-representable* if there exists an injective morphism that maps it into the multiplicative monoid of square matrices of dimension k where the entries belong to a semiring (typically $\mathbb{N}$ or $\mathbb{C}$, but this will be specified if necessary). We may also drop the integer k and say it is representable to mean that it is k-representable for some k. The injective morphism is also called a *faithful representation* of the monoid. Here we investigate this question in the particular case of trace monoids. It is fairly simple to answer this question by resorting to standard existing constructions. It is less clear how to compute the minimum dimension k for a given trace monoid. We will show that the family of 2-representable trace monoids can be characterized and this is the main result of this work.

2.2 Representation of trace monoids by matrices

It is well know that every free monoid can be faithfully represented by matrices with entries in $\mathbb{N}$, i.e., that there exists an injective morphism of the free monoid A^* into the multiplicative monoid of 2×2-matrices with entries in $\mathbb{N}$. If the free monoid has two generators, the two representations to be found in the literature are defined by the following images X and Y of the two generators (cf. e.g., [7, p.156]):

$$X = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \quad Y = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \tag{1}$$

and

$$X = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \quad Y = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}$$

It is clear that the first two matrices generate a free monoid since to every vector $(a\ b)^\top$ with $a > 0$ and $b > 0$, X assigns a vector $(c\ d)^\top$ where $c < d$, while Y assigns a vector $(c\ d)^\top$ where $c > d$. Concerning the last two matrices, compute the matrix $Z_1 \ldots Z_n$ where $Z_i = X$ or Y for $i = 1, \ldots, n$. One verifies easily that it is equal to:

$$\begin{pmatrix} 2^n & 0 \\ N & 1 \end{pmatrix}$$

where N is the integer whose binary expansion is $Z_n \ldots Z_1$ when X is interpreted as 0 and Y as 1 (with possibly leading 0's). This representation can be extended in two steps. First we may substitute any integer k for 2, obtaining thus a faithful representation of a free monoid with k generators. Second, any finite direct product of free finitely generated monoids can be faithfully represented. E.g., in the case of the direct product of the free monoid generated by $a_1, \ldots, a_m$ and the free monoid generated by $b_1, \ldots, b_p$ it suffices to consider the morphism, [12]

$$\phi(a_i) = \begin{pmatrix} m & 0 & 0 \\ 0 & 1 & 0 \\ i & 0 & 1 \end{pmatrix} \quad \phi(b_j) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & p & 0 \\ 0 & j & 1 \end{pmatrix}$$

More generally, the direct product of k free monoids is $k + 1$ representable. This has a direct consequence on the question under study. Indeed, it is well-known that every trace monoid can be injectively mapped into a direct product of free monoids. Furthermore, the number of free monoids in such a direct product is less than or equal to the number of cliques of I (as graph with the letters of the alphabet as vertices) whose union equals the dependence relation. Then the previous observation yields the following

Proposition 1 *For every trace monoid M there exists in injective morphism ϕ of M into $\mathbb{N}^{(k+1)\times(k+1)}$ where k equals the minimum number of cliques covering the graph of the dependence relation.*

2.3 Decision issues

The decision problem is almost hopeless since it has been proven that given two 3×3 integer matrices it is undecidable to determine whether or not they generate a free monoid, cf. [12], let alone a given trace monoid. This result was strenghened upper triangular matrices, [2]

Proposition 2 *Given a finite set of 3×3 integer matrices, it is undecidable whether or not they generate a free monoid.*

Proof. Indeed, by [5] it is undecidable whether or not a finite subset of the direct product $A^* \times A^*$ where A contains at least two elements, is a code, i.e., a free submonoid. ■

Worse, the special case of the square matrices of dimension 2 is still open and it is given some evidence in [2] why the answer to this question is probably not obvious. As a striking example, the authors leave unanswered the question whether or not the following two matrices generate a free monoid

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 0 & 5 \end{pmatrix}$$

The commutative case is simpler. Actually, we have the following

Proposition 3 *Given $X, Y \in \mathbb{N}^{2\times 2}$ it is recursively decidable whether or not they generate a free commutative submonoid.*

Proof. Indeed, assume first X and Y commute which can be easily. The submonoid they generate is not free if and only if the matrices X and Y satisfy an equality of the form

$$X^n = Y^m \tag{2}$$

for some $n, m > 0$. Observing that X and Y generate a free commutative monoid if and only if so do X^2 and Y^2, we may assume without loss of generality that the two matrices have nonnegative determinants. In particular their determinants are both equal to 1 or both equal to 0 else both different from 1 and from 0.

Consider first the case where the determinants are both equal to 0. If further X (or Y) has no nonzero eigenvalue then $X^2 = 0$. Thus from now on X and Y have different eigenvalues and since they are simultanously diagonalizable there exists some non-singular matrix P such that

$$P^{-1}XP = \begin{pmatrix} a/2 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and } P^{-1}YP = \begin{pmatrix} b/2 & 0 \\ 0 & 0 \end{pmatrix} \tag{3}$$

where a and b are the traces of X and Y. Condition (2) is equivalent to $(\frac{a}{2})^n = (\frac{b}{2})^m$ which can be easily checked for.

Consider now $\mathrm{Det}(X) = \mathrm{Det}(Y) = 1$. We can verify that in this case, the submonoid is never free. Since the monoid of matrices with determinant 1 is freely generated by the two matrices of (1), the condition $XY = YX$ can be interpreted as an equation in the free 2-generator monoid. It is well-known that this equation implies that there exist two integers p and r and a matrix Z which is a product of occurrences of the matrices L and R such that $X = Z^p$ and $Y = Z^r$, cf., e.g., [4, Theorem 4. 2. p. 359]. This yields $X^r = Y^p$.

Now assume the two determinants are different from 0 and 1. In particular, equality (2) implies

$$\mathrm{Det}(X)^n = \mathrm{Det}(Y)^m$$

for some integers n, m. In fact, given the two matrices, the decomposition of the two determinants in prime factors can be readily computed and thus the existence of n and m can be checked easily. More precisely, if such integers exist then all the possible pairs (n, m) are of the form (kn_0, km_0) where n_0 and m_0 are fixed integers that are computable from the determinants and k is an arbitrary integer. Let $\lambda_1 \geq \lambda_2 > 0$ and $\mu_1 \geq \mu_2 > 0$ be the two roots of the characteristic polynomial of X and Y respectively. Then (2) implies $\lambda_1^n = \mu_1^m$ and $\lambda_2^n = \mu_2^m$, i.e, $\mathrm{Det}(X)^n = (\lambda_1\lambda_2)^n = (\mu_1\mu_2)^m = \mathrm{Det}(Y)^m$. This means with the previous notations, $n = kn_0$ and $m = km_0$. By extracting the k-th real positive root we get

$$\lambda_1^{n_0} = \mu_1^{m_0} \text{ and } \lambda_2^{n_0} = \mu_2^{m_0} \tag{4}$$

Now this is an equality between two algebraic integers and it can thus be tested in polynomial time, [10]. Conversely, assume equality (4) holds. If the roots are different, then it implies that the two matrices X^{n_0} and Y^{m_0} are equal. It remains the case where both roots have multiplicity 2, i.e., $\lambda_1 = \lambda_2$ and $\mu_1 = \mu_2$. An easy computation shows that the matrices (or their transposed) are of the form

$$X = \begin{pmatrix} a & c \\ 0 & a \end{pmatrix} \; Y = \begin{pmatrix} b & d \\ 0 & b \end{pmatrix}$$

where $a, b, c, d \in \mathbb{N}$. For some k with $n = kn_0$, $m = km_0$, equality (2) reads $a^n = b^m$ and $na^{n-1}c = mb^{m-1}d$. This latter condition, when the former is satisfied, is equivalent to $ncd - mab = 0$. By dividing by k we obtain $a^{n_0} = b^{m_0}$ and $n_0cd - m_0ab = 0$ which can be trivially verified.

■

3 Free products of free commutative monoids

This section is devoted to the proof of the main result of this work. In one direction it is an immediate consequence of the results of Proposition (4). In the other direction it relies on a specific construction on matrices heavily based on the uniqueness of the decomposition of integers into prime factors.

We start with an elementary result on commuting matrices which will yield the necessity of the condition for a trace monoid to be 2-representable. As usual we denote by I the identity matrix (at the risk of some possible confusion with the independence relation, but we prefer to keep the usual terminology) and the complex field by $\mathbb{C}$. More precisely, we investigate the case of 2×2-matrices over the complex numbers. This is simpler than the general case, see [8] and can be treated directly.

Proposition 4 *Let X, Y, Z be three matrices in $\mathbb{C}^{2\times 2}$ satisfying $XY = YX$, $XZ = ZX$ and $YZ \neq ZY$. Then X is a scalar matrix.*

Proof. This is an immediate consequence of the following claim:

Let X, Y be two matrices in $\mathbb{C}^{2\times 2}$ and assume X is non-scalar. Then $XY = YX$ if and only if Y is of the form $Y = \lambda I + \nu X$ where $\lambda, \nu \in \mathbb{C}$.

Indeed, assume first X has two distinct characteristic roots. For some matrix P, we have $X' = P^{-1}XP$ is diagonal. Then the matrix $Y' = P^{-1}YP$

commutes with X' and it is easy to check that it is a linear combination of X' and I and Y is itself a linear combination of X and I. Now if the matrix X has a unique eigenvalue, then for some matrix P, we have $X' = P^{-1}XP$ is of the form

$$\begin{pmatrix} a & d \\ 0 & a \end{pmatrix}$$

It is then trivial to check that the matrix $Y' = P^{-1}YP$ commutes with X' and is of the same form. Then Y' is again a linear combination of X' and I and thus Y is itself a linear combination of X and I.

■

We are now ready to tackle the main result.

Theorem 1 *If a trace monoid has a faithful representation by 2×2-matrices with entries in $\mathbb{C}$ then it is the direct product of a free commutative monoid with the free product of free commutative monoids.*

Conversely, if it is the direct product of a free commutative monoid with the free product of free commutative monoids then it can be faithfully represented by 2×2-matrices with entries in the natural numbers $\mathbb{N}$

Proof. Let $M(A, I)$ be a 2-representable trace monoid. We will show that A can be decomposed into B and C in such a way that:

1) the submonoid generated by B is a free commutative monoid

2) the submonoid generated by C is the free product of free commutative monoids

3) $M(A, I)$ is the direct product of the submonoids generated by B and C.

Let B be the subset of all elements of A commuting with every element in A. It suffices to verify that the submonoid generated by $C = A - B$ is a free product of free commutative monoids. Indeed, if this were not true then there would exist three letters $x, y, z \in C$ such that: $xy \sim yx$, $yz \sim zy$ and $zx \nsim xz$. By Proposition (4) this implies that the matrix associated with y is scalar, thus commutes with all matrices, a contradiction.

To show the converse, we start with the special case of the free product of an arbitrary number of free commutative monoids for which we exhibit a particular representation. We first observe that two matrices:

$$\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \text{ and } \begin{pmatrix} c & d \\ 0 & 1 \end{pmatrix}$$

commute iff $(a-1)d = (c-1)b$. Denoting by $r = b/(a-1) = d/(c-1)$ the common value, each matrix belonging to the submonoid generated by these two commuting matrices is of the form:

$$\begin{pmatrix} x & y \\ 0 & 1 \end{pmatrix} \tag{5}$$

with $r = y/(x-1)$ i.e. $y = rx - r$.

Let $A_1, A_2, \ldots, A_n$ be the disjoint subalphabets such that the trace monoid is the free product of the free commutative monoids generated by each of them and let A be the union of these subalphabets. Set $m = \max\{n, \max\{\mathrm{Card}(A_i) : i = 1, \ldots n\}\}$ and choose m different prime integers $r_j, j = 1, \ldots n$ such that no prime r_j divides the difference between two different r_k's. With each element $a \in A$ associate an integer of the form:

$$k_a = r_1^{\alpha_1} r_2^{\alpha_2} \ldots r_m^{\alpha_m} \tag{6}$$

where $\alpha_i > 0$ is a positive integer for $i = 1, \ldots, m$ in such a way that the vectors $(\alpha_1, \alpha_2, \ldots, \alpha_m)$ associated with symbols of the same subalphabet A_i are linearly independent. The representation of the trace monoid which we propose is defined by associating with each letter $a \in A_i$ the matrix:

$$\begin{pmatrix} k_a & r_i k_a - r_i \\ 0 & 1 \end{pmatrix} \tag{7}$$

In particular the matrices associated with all letters of the subalphabet A_i generate a free commutative monoid. We shall verify that every matrix different from the identity matrix belonging to the monoid generated by the matrices (7) can be uniquely factorized as a product:

$$X_1 X_2 \ldots X_p \tag{8}$$

where for each $\nu = 1, \ldots, p$, X_ν is different from I and is a product of matrices associated with the same subalphabet A_{j_ν}, while two consecutive factors X_ν and $X_{\nu+1}$ are associated with different subalphabets. Thus, X_ν is as follows:

$$X_\nu = \begin{pmatrix} k_\nu & r_{j_\nu} k_\nu - r_{j_\nu} \\ 0 & 1 \end{pmatrix} \tag{9}$$

Identify the matrix (5) with the linear transformation of the projective line: $z \to \lambda z + \mu$ with

$$\lambda = k_p k_{p-1} \dots k_1$$

and

$$\mu = k_p k_{p-1} \dots k_2 r_{j_1}(k_1 - 1) + k_p k_{p-1} \dots k_3 r_{j_2}(k_2 - 1) + \dots + k_p r_{j_{p-1}}(k_{p-1} - 1) + r_{j_p}(k_p - 1)$$

Assume further λ and μ can be obtained in two different ways, i.e., there exist integers $q \geq p$, $\ell_1, \ell_2, \dots, \ell_q$ and a sequence $1 \leq i_1, i_2, \dots, i_q \leq m$ such that

$$k_p k_{p-1} \dots k_1 = \ell_q \ell_{q-1} \dots \ell_1 \tag{10}$$

and

$$\begin{aligned} &k_p k_{p-1} \dots k_2 r_{j_1}(k_1 - 1) + \dots + k_p r_{j_{p-1}}(k_{p-1} - 1) + r_{j_p}(k_p - 1) \\ &= \ell_q \ell_{q-1} \dots \ell_2 r_{i_1}(\ell_1 - 1) + \dots + \ell_q r_{i_{q-1}}(\ell_{q-1} - 1) + r_{i_q}(\ell_q - 1) \end{aligned} \tag{11}$$

Also, we assume p is the minimum value for which the conditions (10) and (11) hold with $q \neq p$. Because k_p and ℓ_q are divisible by $r_1, \dots, r_m$, we get $r_{j_p} = r_{i_q}$ and therefore

$$\begin{aligned} &k_p k_{p-1} \dots k_2 r_{j_1}(k_1 - 1) + \dots + k_p r_{j_{p-1}}(k_{p-1} - 1) + r_{j_p} k_p \\ &= \ell_q \ell_{q-1} \dots \ell_2 r_{i_1}(\ell_1 - 1) + \dots + \ell_q r_{i_{q-1}}(\ell_{q-1} - 1) + r_{i_q} \ell_q \end{aligned}$$

Now assume $k_p \neq \ell_q$, say for some $1 \leq i \leq m$ and $s > 0$, k_p is divisible by r_i^s but not by r_i^{s+1} and ℓ_q is divisible by r_i^{s+1}. Since all $k_{p-1}, \dots, k_2, k_1$ are divisible by r_i, the sum of the last two terms of the left handside is divisible by r_i^{s+1}, i.e., $r_{j_{p-1}} k_{p-1} - r_{j_{p-1}} + r_{j_p}$ is divisible by r_i. This implies that $-r_{j_{p-1}} + r_{j_p}$ is itself divisible by r_i contradicting the assumptions on the choice of the r's. Thus, equations (10) and (11) hold with $p-1$ and $q-1$ instead of p and q respectively, contradicting the minimality of p and proving the unique factorization.

To complete the proof, we observe the following. Let M_1 be a free product of free commutative monoids, let A be its set of generators and let $\phi : M_1 \to \mathbb{N}^{2\times 2}$ be an injective morphism as just described. Let M_2 be a free commutative monoid and let $p_1, \ldots, p_s$ be prime numbers in one-to-one correspondence with the generators b_i for all $i = 1, \ldots, s$ of M_2. We assume that together with the primes r_i's used in defining ϕ, all these primes are different from one another. Since the determinants of all $\phi(x)$'s with $x \in M_1$ admit only the r_i's as prime factors, it is clear that the morphism $\psi : M_1 \times M_2 \to \mathbb{N}^{2\times 2}$ defined by $\psi(a, b_i) = p_i\phi(a)$ is injective. ■

4 Conclusion

Because of paragraph (2.2) all trace monoids whose dependence graph can be covered by at most k different cliques have a faithful representation by $(k+1)\times(k+1)$ integer matrices. Can this upper bound be lowered? A stronger claim would assert a uniform bound on the dimension of the matrices, which seems quite unlikely.

References

[1] M. Albert and J. Lawrence. A proof of Ehrenfeucht's conjecture. *Theoret. Comput. Sci.*, 41:121–123, 1985.

[2] J. Cassaigne, T. Harju, and J. Karhumäki. On the undecidability of freeness of matrix semigroups. *Internat. J. Algebra Comput.*, 9:295–305, 1999.

[3] C. Choffrut. A remark on the representation of free partially commutative monoids. *Semigroup Forum*, 40:143–152, 1989.

[4] C. Choffrut and J. Karhumäki. Combinatorics of words. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages*, number 1450, pages 329–438. Springer Verlag, 1997.

[5] M. Chrobak and W. Rytter. Unique decipherability for partially commutative alphabets. *Fund. Inform.*, X:323–336, 1987.

[6] M. J. Collins. *Representation and characters of finite groups*, volume 22. Cambridge studies in advanced mathematics, 1990.

[7] S. Eilenberg. *Automata, Languages and Machines*, volume A. Academic Press, 1974.

[8] F.R. Gantmacher. *Matrizenrechnung, Teil I.* VEB Deutscher Verlag der Wissenschaften, 1958.

[9] H. J. Hoogeboom and G. Rozenberg. In V. Diekert and G. Rozenberg, editors, *The Book of Traces*. World Scientific, 1995.

[10] L. Langemyr. Algorithms for a multiple algebraic extension. In C. Traverso T. Mora, editor, *Effective Methods in Algebraic Geometry*, number 94, pages 234–248. Birkhäuser, 1991.

[11] W. Magnus. The use of 2 by 2 matrices in combinatorial group theory. *Resultate der Mathematik*, 4:171–192, 1981.

[12] M. S. Paterson. Unsolvability in 3×3 matrices. *Studies Appl. Math.*, 49:105–107, 1970.