The Cayley-Hamilton Theorem for Noncommutative Semirings

Radu Grosu

Department of Computer Science, Stony Brook University
Stony Brook, NY 11794-4400, USA

Abstract. The Cayley-Hamilton theorem (CHT) is a classic result in linear algebra over fields which states that a matrix satisfies its own characteristic polynomial. CHT has been extended from fields to commutative semirings by Rutherford in 1964. However, to the best of our knowledge, no result is known for noncommutative semirings. This is a serious limitation, as the class of regular languages, with finite automata as their recognizers, is a noncommutative idempotent semiring. In this paper we extend the CHT to noncommutative semirings. We also provide a simpler version of CHT for noncommutative idempotent semirings.

1 Introduction

The continuous dynamics of each mode of a linear hybrid automaton (HA) [2, 7] is a linear system, as is the discrete switching logic among modes [5]. However, the former operates on a vector space, whereas the latter operates on a semimodule [4]. As a consequence, understanding which properties of linear systems hold in both vector spaces and semimodules, and which do not, is essential for developing a formal foundation and associated analysis tools for HAs.

Vector spaces (VSs) have a long history and consequently a large variety of analysis techniques. A defining aspect of a VS is the associated field of scalars, a structure possessing two operations, addition and multiplication, together with their identities and their inverses. Both are associative and commutative and multiplication distributes over addition. For example, $\mathbb{R}$ and $\mathbb{C}$ are fields.

Finding the fixpoint of the equation $x = Ax+B$ in a VS seems to be routine: $x = (I - A)^{-1}B$. However, this is far from being true. For large systems, computing the inverse of a matrix is expensive, and one often uses iterative Gauss-Seidel or Jacobi techniques [8]. Both are based on the identity $(I - A)^{-1} = A^*$, where $A^* = \sum_{n\in\mathbb{N}} A^n$, and converge if the eigenvalues of A are in the unit disc of $\mathbb{C}$.

The above solution for x leads to an amazing conclusion: fixpoint computation does not require inverses or commutativity of multiplication. Hence, one may consider a weaker structure, that lacks subtraction and division, and whose multiplication is not necessarily commutative. Such a structure is called a semiring, and it admits fixpoints of the form $x = a^*b$ for $x = ax + b$. A vector space where the field is replaced with a semiring is called a semimodule.

In general a semiring may admit many fixpoints, and a^*b is the least one. But to single out the least, one needs a partial order, which may be defined canonically as follows: $a \leq b$ if there is a c such that $b = a + c$. This is possible for example in $\mathbb{N}$ but not in $\mathbb{Z}$, as for any $a, b \in \mathbb{Z}$, $a + (-a + b) = b$. Hence, a semiring

cannot have both a canonical order and an inverse. This is where classic and discrete mathematics diverge [4]. So what else does (not) hold in both settings?

In [5] we have shown that minimization of nondeterministic finite automata (NFA) can be advantageously cast as reachability and observability reductions of linear systems. This also allowed us to show that minimal NFA are linear transformations of each other. This result is noteworthy, as no reasonable way of relating minimal NFA was previously known (see for example [1]).

In this paper we continue the above line of work, by investigating the classic Cayley Hamilton theorem (CHT), in the context of noncommutative semirings. The class of regular languages, with NFA as their recognizers, is an important and strongly motivating subclass of these semirings.

For a matrix A with entries in a field, CHT states that A satisfies its own characteristic polynomial, that is $\mathrm{cp}_A(A) = 0$ where $\mathrm{cp}_A(s) = \det(sI - A)$. Hence, any extension of CHT to noncommutative semirings has to solve two orthogonal problems: 1) The lack of subtraction; and 2) The lack of commutativity. They are both critical ingredients in the computation of the determinant $\det(sI - A)$.

The lack of subtraction was addressed in 1964 by Rutherford [10]. The main idea is to define subtraction in terms of addition by replacing terms $a - b$ with pairs (a, b). Consequently, $\det(sI - A)$ becomes $(\det^+(sI - A), \det^-(sI - A))$, a bideterminant, and CHT becomes $\mathrm{cp}^+_A(A) = \mathrm{cp}^-_A(A)$. This allowed Rutherford to extend CHT to matrices with entries in a commutative semiring.

The lack of commutativity is addressed in this paper. The main idea is to define a commutative multiplication in terms of multiplication and addition by replacing products ab with their permutation closure $[\![ab]\!] = ab + ba$. Consequently, $\det(sI - A)$ becomes $[\![\det(sI - A)]\!]$, what we call a pideterminant, and CHT becomes $[\![\mathrm{cp}_A(A)]\!] = 0$. This allows us to extend CHT to any noncommutative structure, and in particular to noncommutative rings.

Combining Rutherford's solution with our own solution, allows us to extend CHT to noncommutative semirings as $[\![\mathrm{cp}^+_A(A)]\!] = [\![\mathrm{cp}^-_A(A)]\!]$. We argue that both solutions are also canonical, in the sense that $\det(A) = \det^+(A) - \det^-(A)$ and that $\det(A) = (1/n!)[\![\det(A)]\!]$ in any field.

Interpreting matrix A as a process, we also observe that the power A^n occurring in CHT can be understood in two ways: 1) As n copies of process A that interleave a single move; 2) As one copy of process A that performs n moves. This observation gives a computational justification for permutation closure, and paves the way to two different forms of CHT. Finally, considering addition idempotent, as in the class of languages, leads to a simpler form of CHT.

The rest of the paper is organized as follows. Sections 2 and 3 review semirings, fields and permutations. Section 4 reviews (bi)determinants, and introduces our new notion of pideterminant. Section 5 follows the same pattern for characteristic (bi)polynomials and pipolynomials. Sections 6 and 7 extend CHT by allowing or disallowing interleaving, respectively, and prove the validity of these extensions in noncommutative semirings. Section 8 particularizes the second version of CHT to idempotent noncommutative semirings. Finally, Section 9 contains our concluding remarks and directions for future work.

2 Semirings and Fields

A *semiring* $\mathbb{S} = (S, +, \times)$ is a set S possessing two binary operations, *addition* and *multiplication* together with their *identities* 0 and 1, respectively. Addition is required to be *associative and commutative* and multiplication is required to be *associative*. Multiplication *distributes* over addition and 0 is an *absorbing* element for multiplication, that is, for all $a \in S$, $a \times 0 = 0 \times a = 0$.

A semiring where multiplication is commutative is called a *commutative semiring*, and a semiring where addition is idempotent (for all $a \in S$, $a + a = a$) is called an *idempotent semiring*. For example, the natural numbers $\mathbb{N}$ with the usual meaning of $+$ and $\times$, form a commutative semiring.

The class of *languages* over a finite alphabet A is an idempotent semiring $\mathbb{A} = (\wp(A^*), +, \times)$, where the elements in A^* are *words* (sequences) over A, the elements in $\wp(A^*)$ are sets of words (languages), $+$ is interpreted as *union* and $\times$ is interpreted as *concatenation*. A language is *regular*, if it is accepted by a finite automaton. For example, $\{a\}$ and $\{b\} \sum_{n \in \mathbb{N}} \{a\}^n$ are regular languages.

A *field* $\mathbb{F} = (F, +, \times)$ is a commutative semiring where $+$ and $\times$ have inverses. For example, $\mathbb{R}$ and $\mathbb{C}$ are fields, with the usual meaning of $+$ and $\times$.

3 Permutations

A *permutation* π is a bijection of the finite set $\{1,\ldots,n\}$ onto itself. It has associated a *directed graph* $G(\pi) = (V, E)$, with set of vertices $V = \{1 \ldots n\}$, and set of edges E, consisting of pairs $(i, \pi(i))$, one for each for $i \in V$. For example:

$$\pi = \{(1,2), (2,3), (3,4), (4,1), (5,7), (6,6), (7,5)\}$$

is a permutation of the set $\{1,\ldots,7\}$. Its associated graph $G(\pi)$, which is shown in Figure 1, illustrates an important property of the graph of any permutation: it decomposes into *elementary disjoint cycles*, the *partial rotations* of π.

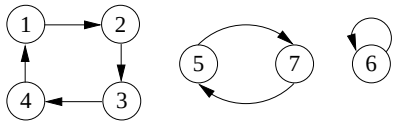

Fig. 1. The graph $G(\pi)$ of permutation π.

If $G(\pi)$ has an even (odd) number of cycles with even number of edges, then π is said to have positive (negative) *sign*. For example, the graph $G(\pi)$ above, has two cycles of even length, so π has positive sign. Denote by $P(n)$ be the set of all permutations of the set $\{1 \ldots n\}$, and by $P^+(n)$ and $P^-(n)$, its positive and negative subsets, respectively.

A *partial permutation* π of the finite set $V = \{1,\ldots,n\}$ is a permutation of a subset S of V. For example, if $V = \{1 \ldots 7\}$ and π is defined as follows:

$$\pi = \{(1,2), (2,1), (3,4), (4,6), (6,3)\}$$

then π is a permutation of its domain $\mathrm{dom}(\pi) = \{1, 2, 3, 4, 6\}$, and a partial permutation of V. Every partial permutation π of V can be extended to a permutation $\hat{\pi}$ of V by letting $\hat{\pi}(i) = \pi(i)$ if $i \in \mathrm{dom}(\pi)$ and $\hat{\pi}(i) = i$, otherwise.

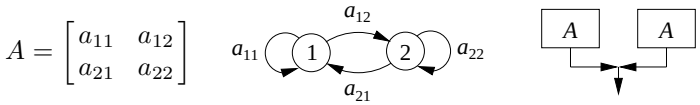

Fig. 2. (a) Matrix A. (b) $G(A)$. (c) $[\![A^2]\!]$.

Given a permutation $\pi \in P(n)$, we write $\overline{\pi}$ for the sequence $(1, \pi(1)) \ldots (n, \pi(n))$. We extend permutations $\pi \in P(n)$ to sequences $w = w_1 \ldots w_n$ in a componentwise fashion: $\pi(w) = w_{\pi(1)} \ldots w_{\pi(n)}$. For example, if $w = abcdefg$ and π is the permutation shown in Figure 1, then $\pi(w) = dabcgfe$. Similarly, if σ is another permutation of the set $\{1, \ldots 7\}$ then $\pi(\overline{\sigma})$ is equal to:

$$(4, \sigma(4))(1, \sigma(1))(2, \sigma(2))(3, \sigma(3))(7, \sigma(7))(6, \sigma(6))(5, \sigma(5))$$

4 The Determinant in Noncommutative Semirings

A square matrix A of order n with entries in a field $\mathbb{F}$ is an element of $F^{n \times n}$. One says that A has n rows and n columns. For example, a matrix A of order 2 (a 2 by 2 matrix) is shown in Figure 2(a). Row 1 is (a_{11}, a_{12}), row 2 is (a_{21}, a_{22}), column 1 is (a_{11}, a_{21}) and column 2 is (a_{21}, a_{22}).

A square matrix A of order n has associated a *weighted directed graph* $G(A) = (V, E, A)$, where $V = \{1,\ldots,n\}$ is the set of *vertices* and $E = \{(i, j) \in V^2 \mid A_{ij} \neq 0\}$ is the set of *edges* (i, j) with *weight* A_{ij}. For example, the graph $G(A)$ of the above matrix A of order 2 is shown in Figure 2(b).

A *generalized path* p in $G(A)$ is a sequence of edges $p_1 \ldots p_n$ in E. This is called a *path* if $\mathrm{head}(p_i) = \mathrm{tail}(p_{i+1})$ for each $i < n$.[1] The product $p(A) = A_{p_1} \ldots A_{p_n}$ is called the *weight of* p. For example, $(1, 1)(1, 2)(2, 1)(A) = A_{11}A_{12}A_{21}$. A path that starts and ends with same vertex is called a *cycle*. This is called *simple* if it has no other repeated vertices.

Using permutations, generalized paths and associated path weights, one can explicitly define the *determinant* of a square matrix A of order n, as follows:

$$\det(A) \;=\; \left(\textstyle\sum_{\pi \in P^+(n)} \overline{\pi} \;-\; \sum_{\pi \in P^-(n)} \overline{\pi}\right)(A)$$

where each term of det is applied to A. We denote by $\det^+$ and $\det^-$ the positive and the negative parts of the determinant operator det, respectively.

Since the determinant is an n-linear function, its value is typically computed iteratively, by expanding it along one of the rows (or columns) of its argument matrix, and then repeating the process for each remaining submatrix, until the argument matrix has only one entry (Laplace expansion).

Rutherford has transfered the determinant's computation from a commutative semirings to a ring-like structure by defining subtraction in terms of addition of tuples. Hence, the determinant has become a tuple, called a *bideterminant*:

$$\mathrm{bdt}(A) = (\det^+,\ \det^-)(A) = (\det^+(A),\ \det^-(A))$$

The bideterminant can be computed by linear expansion, as discussed above, by pretending first that negation was available to compute $\det(A)$, and then separating the positive and the negative parts of the result.

[1] This definition of paths is more convenient in our setting.

For example, consider the matrix A of Figure 2(a). The set $P(2)$ has only two permutations, which are also rotations: π_1 and π_2:

$$\pi_1 = \{(1,1),\,(2,2)\} \in P^+(2), \quad \pi_2 = \{(1,2),\,(2,1)\} \in P^-(2)$$

Using the Laplace expansion, first for row 1 and then for row 2 of A, and first for row 2 and then for row 1, one obtains the following bideterminants:

$$\mathrm{bdt}_{12}(A) = (a_{11}a_{22},\, a_{12}a_{21}), \quad \mathrm{bdt}_{21}(A) = (a_{22}a_{11},\, a_{21}a_{12})$$

In commutative semirings $\mathrm{bdt}_{12}(A) = \mathrm{bdt}_{21}(A)$. In *noncommutative semirings*, however, this is generally not true, i.e., $\mathrm{bdt}_{12}(A) \neq \mathrm{bdt}_{21}(A)$.

For example, in regular languages, the graph $G(A)$ corresponds to a *finite automaton*, and F is a finite set, called the *input alphabet*. As a consequence, the sequence (word) of inputs $a_{11}a_{22}$ is different from $a_{22}a_{11}$, unless $a_{11} = a_{22}$.

While extensive work has been devoted to determinants of matrices with entries in noncommutative rings [3], the author is not aware of any definition of determinants for matrices with entries in noncommutative semirings. Moreover, the definitions of determinants in noncommutative rings, for example the *quasideterminants* of [3], do not have determinants as a particular case, and involve division. This operation, however, is not available in semirings.

Inspired by Rutherford, we transfer the determinant's computation to a structure possesing a commutative multiplication defined in terms of addition and multiplication. This is equivalent to extending the notion of determinant to a *pideterminant* which is the sum of all row (or column) expansions. Hence:

$$\mathrm{pdt}(A) = (a_{11}a_{22} + a_{22}a_{11},\, a_{12}a_{21} + a_{21}a_{12})$$

Note that if the semiring is commutative, $\mathrm{pdt}(A) = 2!\,\mathrm{bdt}(A)$. Let $\pi(\mathrm{bdt}_{12})$ be defined as $(\pi(\mathrm{bdt}^+_{12}), \pi(\mathrm{bdt}^-_{12}))$, the π-permutation of bdt_{12}. Then:

$$\mathrm{bdt}_{12}(A) = \pi_1(\mathrm{bdt}_{12})(A), \qquad \mathrm{bdt}_{21}(A) = \pi_2(\mathrm{bdt}_{12})(A)$$

In general, the expansion of a determinant for rows $r_1, \ldots, r_n$, in this order, results in a permutation π of $\mathrm{bdt}_{12\ldots n}$, where $\pi = \{(1,\, r_1), \ldots, (n,\, r_n)\}$. Moreover, expanding recursively all rows of a matrix results in all possible permutations, and the positive (negative) sign of the arguments is preserved.

Hence, given a matrix A of order n with entries in a noncommutative semiring, one obtains the following *explicit* representation of a pideterminant:

$$\mathrm{pdt}\,(A) \;=\; \Big(\sum_{\substack{\pi \in P^+(n)\\ \sigma \in P(n)}} \sigma(\overline{\pi}),\;\; \sum_{\substack{\pi \in P^-(n)\\ \sigma \in P(n)}} \sigma(\overline{\pi})\Big)\,(A)$$

For notational simplicity we denote the *permutation closure* $\sum_{\pi \in P(n)} \pi(w)$ of w as $[\![\, w \,]\!]$. Using this notation we can write $\mathrm{pdt} = ([\![\,\det^+]\!],\, [\![\det^-]\!])$. To further simplify the notation we will write when convenient $[\![\,\det^+(A)\,]\!]$ for $[\![\,\det^+]\!](A)$, and pretend that we worked in a free (permutation closed) semiring.

5 The Characteristic Polynomial in Noncomm. Semirings

The *characteristic polynomial* $\mathrm{cp}_A(s)$ in indeterminate s, associated to a matrix A of order n with entries in a field, is defined as $\det(sI - A)$. For example, for the second order matrix A of Figure 2(a) one has:

$$\mathrm{cp}_A(s) \;=\; \det\left(\begin{bmatrix} s - a_{11} & -a_{12} \\ -a_{21} & s - a_{22} \end{bmatrix}\right) \;=\; s^2 - (a_{11} + a_{22})s + (a_{11}a_{22} - a_{12}a_{21})$$

The characteristic polynomial is used to compute the eigenvalues of a matrix A with entries in a real field by finding the roots of the equation $\mathrm{cp}_A(s) = 0$. Eigenvalues and their associated eigenvectors are essential tools for computing the explicit solution of systems of linear difference and differential equations.

The characteristic polynomial was generalized by Rutherford to a *characteristic bipolynomial* $\mathrm{cbp}_A(s)$ for matrices with entries in a commutative semiring:

$$\mathrm{cbp}_A(s) = (\mathrm{cp}_A^+(s),\ \mathrm{cp}_A^-(s))$$

This polynomial can be computed by first pretending one works in a field, and then separating the positive and the negative terms. For matrix A of Figure 2(a):

$$\mathrm{cbp}_A(s) = (s^2 + a_{11}a_{22},\ (a_{11} + a_{22})s + a_{12}a_{21})$$

We define the *characteristic pipolynomial* $\mathrm{cpp}_A(s)$ of a matrix A with entries in a noncommutative semiring as follows:

$$\mathrm{cpp}_A(s) = (\llbracket \mathrm{cp}_A^+(s) \rrbracket,\ \llbracket \mathrm{cp}_A^-(s) \rrbracket) = (\mathrm{cpp}_A^+(s), \mathrm{cpp}_A^-(s))$$

To compute $\mathrm{cpp}_A(s)$ one can pretend to work in a free semiring when computing the closure of $\mathrm{cbp}_A(s)$. For example, for matrix A matrix of Figure 2(a):

$$\mathrm{cpp}_A(s) = (\llbracket s^2 + a_{11}a_{22} \rrbracket,\ \llbracket (a_{11} + a_{22})s + a_{12}a_{21} \rrbracket)$$

6 Multi-Process CHT for Noncommutative Semirings

The Cayley-Hamilton theorem (CHT) is a classic result in linear algebra over fields stating that a matrix satisfies its own characteristic polynomial: $\mathrm{cp}_A(A) = 0$.

One of the applications of CHT is to compute the dimension of the A-cyclic vector space $V_A = \{A^n \mid n \in N\}$. This vector space is fundamental in the study of observability and controllability of linear systems.

For example, if the state-space description of a linear system is given by the following difference equations:

$$x(n+1) = Ax(n) + Bu(n), \quad y(n) = Cx(n), \quad x(0) = x_0$$

and V_A has dimension k, then the observability and controllability matrices of the system are defined as follows: $O = [C\, CA \ldots CA^{k-1}]^t$, $K = [B\, AB \ldots A^{k-1}B]$.

In [5] we have shown that these matrices are also relevant in the minimization of nondeterministic finite automata (NFA). Moreover, we have proved that minimal NFA are linear transformations of each other. Relating minimal NFAs is a problem that was addressed, but not properly solved, before (see e.g. [1]).

Since $\mathrm{cp}_A(A)$ is a matrix equation, the following *conventions* are used: s is replaced with A, as is replaced with aA and every constant k is replaced with kA^0. For example, for the matrix A of Figure 2(a) one obtains:

$$\mathrm{cp}_A(A) = A^2 - (a_{11} + a_{22})A + (a_{11}a_{22} - \mathtt{a}_{12}a_{21})I = 0$$

This result has been extended to commutative semirings by Rutherford in [10], and a combinatorial proof was given later by Straubing in [11]. The generalized Cayley-Hamilton theorem states that: $\mathrm{cbp}_A^+(A) = \mathrm{cbp}_A^-(A)$.

It is easy to show that CHT does not hold in noncommutative semirings. For example, consider the matrix A of Figure 2(a). Then one would require that:

$$A^2 + a_{11}a_{22}I \;=\; (a_{11}+a_{22})A + a_{12}a_{21}I$$

Now compute the LHS and the RHS of the above equation:

$$\text{LHS} = \begin{bmatrix} a_{11}a_{11}+a_{12}a_{21}+a_{11}a_{22} & a_{11}a_{12}+a_{12}a_{22} \\ a_{21}a_{11}+a_{22}a_{21} & a_{21}a_{12}+a_{22}a_{22}+a_{11}a_{22} \end{bmatrix}$$

$$\text{RHS} = \begin{bmatrix} a_{11}a_{11}+a_{12}a_{21}+a_{22}a_{11} & a_{11}a_{12}+a_{22}a_{12} \\ a_{11}a_{21}+a_{22}a_{21} & a_{12}a_{21}+a_{22}a_{22}+a_{11}a_{22} \end{bmatrix}$$

They are obviously different because of the lack of commutativity of the entries in A. One of the main results of this paper is that a matrix A satisfies its own characteristic pipolynomial: $\text{cpp}^+_A(A) = \text{cpp}^-_A(A)$.

Theorem 1. (MULTI-PROCESS CHT FOR NONCOMMUTATIVE SEMIRINGS) *A matrix A with entries in a noncommutative (semi)ring satisfies its own characteristic pipolynomial. That is,* $\text{cpp}^+_A(A) = \text{cpp}^-_A(A)$.

Consider the CHT equation of the example above. Two offending weights are $a_{11}a_{22}$ of $(\text{LHS})_{11}$ and $a_{22}a_{11}$ of $(\text{RHS})_{11}$. These weights may be not equal in a noncommutative semiring. However, their permutation closure is the same:

$$[\![\, a_{11}a_{22} \,]\!] = [\![\, a_{22}a_{11} \,]\!] = a_{11}a_{22} + a_{22}a_{11}$$

CHT can be intuitively understood as an *incremental construction* of $[\![\text{cbp}^+_A(A)]\!]$ and $[\![\text{cbp}^-_A(A)]\!]$ such that at the end $[\![\text{cbp}^+_A(A)]\!] = [\![\text{cbp}^-_A(A)]\!]$. The construction is justified with the help of the graph $G(A)$ associated with A.

For illustration purpose, we will use the matrix A of order 2 shown in Figure 2 with correponding CHT $[\![\, A^2 + a_{11}a_{22}I \,]\!] = [\![\, (a_{11}+a_{22})A + a_{12}a_{21}I \,]\!]$.

Start with $\text{LHS}_0 = A^2$. Each entry $(A^2)_{ij}$ is a sum of weights of length 2, each weight being associated to a path from i to j in $G(A)$. For example, the path (1,2)(2,1) has associated the weight $a_{12}a_{21}$. Since $G(A)$ has only 2 vertices, all these paths must have at least one simple cycle.

Suppose we first want to add to RHS the weights of the simple cycles of length 2 in $G(A)$. For example, (1,2)(2,1) and (2,1)(1,2), with the associated weights $a_{12}a_{21}$ and $a_{21}a_{12}$, respectively. They are all contained in the *diagonal* $\text{diag}(A^2)$ of A^2. However, since the diagonal of A^2 may also contain products of cycles with length less then 2, we denote by $\text{diag}_s(A^2)$ the *restriction* of $\text{diag}(A^2)$ to simple cycles only. Similarly, we denote by $\text{trace}_s(A^2)$, *the sum* of the simple cycles in $\text{diag}_s(A^2)$. Consequently, we start with: $\text{RHS}_0 = \text{trace}_s(A^2)I$.

All the cycle weights in $\text{trace}_s(A^2)$ are permutations (in fact rotations) of the simple-cycle weights of vertex 1. There are 2 such permutations (including identity) which we denote by π_1 and π_2. Now we can write:

$$\text{RHS}_0 = \textstyle\sum_{i=1}^{2} \pi_i(a_{12}a_{21})I$$

Multiplying these permutations with the identity matrix I has unfortunately undesired consequences: it introduces spurious weights in each entry of RHS_0. For example entry $(\text{RHS}_0)_{11}$ also contains weights such as $a_{21}a_{12}$.

To balance out spurious weights in RHS_0 we have to add the corresponding permutations of A^2 to the LHS. Hence, the LHS becomes $\mathrm{LHS}_1 = \sum_{i=1}^{2} \pi_i(A^2)$. Obviously, this introduces many more spurious weights on the LHS.

The construction now continues by adding and balancing out cycles of length 1 on the RHS, which we omit for space limitations.

Discussion. In the above construction, most of the effort is devoted to *fixing "spurious" weights*. One may therefore wonder whether such weights make any sense, and if not, whether there was a way of getting rid of them.

Consider matrix A of order n and regard $G(A)$ as a process which either acts as an acceptor or as a generator of words over a given alphabet. The power A^n can be interpreted in two distinct ways: 1) As the *interleaving of n copies of A*, each starting in an arbitrary state and performing *one move*; 2) As *one copy of A* that performs *n moves*. In each case, one can ask what is *the sum*, if counting is important, of the words accepted (or generated) by A^n?

In the interleaving interpretation of A^n, cycle weights s.a. $a_{11}a_{22}a_{33}$ make sense: the word is generated by letting the first copy of A start in vertex 1 and make one move to generate a_{11}, then the second start in vertex 2 and make one move to generate a_{22} and finally the third start in vertex 3 and make one move to generate a_{33}. Since every copy of A can make any move before or after the other copy made one move, one has to consider all permutations.

In conclusion, *Theorem 1 explicitly defines the behavior of the process resulting from the interleaving (shuffling) of n copies of process A.* For example, Figure 2(c) shows the interleaving of two copies of matrix A of order 2.

In process algebra, commutativity of inputs is equivalent with their *independence*: one obtains the same result no matter in what order one processes them. Hence, matrices with entries in a commutative semiring correspond to processes over a set of independent inputs. For general processes, independence might hold for some subsets of the input alphabet, but not for the entire alphabet. The knowledge of an independence relation over the input alphabet is still very useful, because it allows to partition matrix A into commutative blocks. Considering only one version of the commutative products, then corresponds to the *partial-order reduction technique* used in computer-aided verification.

7 Single-Process CHT for Noncommutative Semirings

In the second interpretation of A^n in the CHT, as one copy of A performing n moves, "spurious" cycle weights such as $a_{11}a_{22}a_{33}$ or $a_{11}a_{23}a_{32}$ make no sense. We would therefore like to find a way to get rid of them.

An *acceptor algorithm* that cleans up "wrong" weights, is to: 1) First compute $\mathrm{cpp}_A(A)$ as before, and 2) Then remove all generalized path-weights in $\mathrm{cpp}_A(A)_{ij}$ that either do not correspond to paths, or they are are misplaced, that is their corresponding starting vertex is not i or their ending vertex is not j.

For example, the weight of the generalized path $(1,1)(2,2)$ is removed because it is not a path. The wight of $(1,2)(2,1)$ is removed if it appears in $\mathrm{cpp}_A(A)_{22}$.

Theorem 2. (1ST SINGLE-PROCESS CHT FOR NC SEMIRINGS) *A matrix A with entries in a noncommutative semiring satisfies its own characteristic pipolynomial, when clean up is added at the end of the permutation closure.*

One may *avoid introducing "spurious" weights by treating cycles as diagonal matrices*. If p is a cycle, then let $\langle p \rangle$ be the diagonal matrix with $\langle p \rangle_{ii} = p$ if p belongs to $(A^{|p|})_{ii}$ and $\langle p \rangle_{ii} = 0$, otherwise. If c is a sum of cycles of same length, then let $\langle c \rangle$ denotes the sum of their corresponding matrices.

A *generator algorithm* for $\mathrm{cpp}_A(A)$ can now be defined as follows: 1) Take the rotation closure of all cycles. 2) Consider cycles atomic and take the rotation closure of the entire terms. 3) Keep the cycles fixed, and take the partial-permutation closure. We denote by $(\!(\, \mathrm{cbp}_A(A) \,)\!)$ steps 1–3.

For example, let $\langle c \rangle = \langle a_{12}a_{21} + a_{21}a_{12} \rangle$ be the rotation closure of cycle matrix $\langle a_{12}a_{21} \rangle$. Then:

$$(\!(\, \langle a_{12}a_{21} \rangle A^2 \,)\!) = \begin{array}{c} \langle c \rangle A^2 \;+\; A\langle c \rangle A \;+\; A^2\langle c \rangle + \\ \pi_1(\langle c \rangle A^2) + \pi_2(A\langle c \rangle A) + \pi_3(A^2\langle c \rangle) \end{array}$$

where π_1, π_2 and π_3 swap positions 2 and 3, 1 and 3, and 1 and 2, in $\langle c \rangle A^2$ $A\langle c \rangle A$ and $A^2\langle c \rangle$, respectively.

Theorem 3. (2ND SINGLE-PROCESS CHT FOR NONCOMMUTATIVE SEMIRINGS) *In a noncommutative semiring $(\!(\, \mathrm{cbp}^+_A(A) \,)\!) = (\!(\, \mathrm{cbp}^-_A \,)\!)$ for any matrix A if each cycle c is interpreted as the (diagonal) matrix $\langle c \rangle$.*

For example, the CHT for matrix A of order 2 simplifies to:

$$(\!(\, A^2 \,)\!) = \langle a_{11} + a_{22} \rangle A + A\langle a_{11} + a_{22} \rangle + \langle a_{12}a_{21} + a_{21}a_{12} \rangle$$

8 The CHT for Noncommutative Idempotent Semirings

Suppose now that the entries of matrix A belong to a noncommutative *idempotent semiring*. Recall that a semiring S is called idempotent, if its *additive operation is idempotent*, that is, for any element a of the semiring, $a + a = a$. Noncommutative idempotent semirings are important, as they contain as a particular case the class of *regular languages*.

Although counting is not important in such semirings, it is not obvious (at least to the author) how the multi-process CHT could be further simplified. One still needs the permutation closure $[\![\, \mathrm{cbp}_A(A) \,]\!]$, but addition simplifies.

The single-process version of the CHT can be however, further simplified in idempotent semirings. Let us denote by $(\!(\, \mathrm{cbp}_A(A) \,)\!)$ the closure operation discussed in the previous section, where the last step, the partial-permutation closure, is discarded. Then we have the following theorem.

Theorem 4. (SINGLE-PROCESS CHT FOR NC IDEMPOTENT SEMIRINGS) *Let A be a matrix of order n with entries in a noncommutative idempotent semiring. If each cycle c in $\mathrm{cbp}_A(A)$ is interpreted as the (diagonal) matrix $\langle c_k \rangle$, then its Cayley-Hamilton theorem simplifies to $A^n = (\!(\, \mathrm{cbp}^-_A(A) \,)\!)$.*

For example, consider matrix A of Figure 2(a), and assume its entries are distinct and belong to a noncommutative idempotent semiring. Then the CHT of this matrix simplifies to the following form:

$$A^2 = \langle a_{11} + a_{22} \rangle A + A\langle a_{11} + a_{22} \rangle + \langle a_{12}a_{21} + a_{21}a_{12} \rangle$$

9 Conclusions

We have extended the Cayley-Hamilton theorem (CHT), a classic result in linear algebra over fields which states that a matrix satisfies its own characteristic polynomial, to noncommutative semirings.

The pideterminant and the pipolynomial we have defined for this purpose could have also been inferred by using the *shuffle product* of [9] and an evaluation function *eval*. Given a noncommutative ring R one can define a commutative ring $S = (R^*, +, \|)$ where R^* consists of sequences in R and $s\|t$ is the shuffle product of $s, t \in R^*$, defined in terms of addition and concatenation.

Since S is commutative, the computation of $\det(A)$ is well defined and so is the Cayley-Hamilton theorem. As a consequence, $\mathrm{pdt}(A) = eval(\det(A))$ where *eval* replaces concatenation with the product in R and evaluates the result.

In S the power A^n is the n-shuffle (interleaving) of A, and it can be expressed as a linear combination of $I, A, \ldots A^{n-1}$. This observation suggests a generalization of linear dependence for a set of vectors $x_1, \ldots, x_n$ in a noncommutative module as follows: there are scalars $a_1, \ldots, a_n$ such that the shuffle product $a_1\|x_1 + \ldots + a_n\|x_n = 0$. Such extensions are the focus of future work.

References

1. M. Nivat A. Arnold, A. Dicky. A note about minimal nondeterministic automata. *EATCS*, 45:166–169, 1992.
2. R. Alur, C. Courcoubetis, N. Halbwachs, T.A. Henzinger, P.H. Ho, X. Nicolin, A. Olivero, J. Sifakis, and S. Yovine. The algorithmic analysis of hybrid systems. *Theoretical Comp. Sci.*, 138:3–34, 1995.
3. I. Gelfand, S. Gelfand, V. Retakh, and R.L. Wilson. Quasideterminants. *Adv. Math.*, 193:1–82, 2005.
4. M. Gondran and M. Minoux. *Graphs, Dioids and Semirings*. Springer, 2008.
5. R. Grosu. Finite automata as time-invariant linear systems: Observability, reachability and more. In *Proc. of HSCC'09, the 12th International Conference on Hybrid Systems: Computation and Control*, volume 5469 of *LNCS*, pages 194–208, San Francisco, USA, April 2009. Springer Verlag.
6. G.A. Kildall. A unified approach to global program optimization. In *POPL '73: Proceedings of the 1st Annual ACM SIGACT-SIGPLAN Symposium on Principles of Programming Languages*, pages 194–206, New York, NY, USA, 1973. ACM Press.
7. N. Lynch, R. Segala, and F. Vaandrager. Hybrid I/O automata. *Inf. and Comp.*, 185(1):103–157, 2003.
8. W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, New York, NY, USA, 1992.
9. R. Ree. Lie elements and an algebra associated with shuffles. *Anals of Mathematics*, Vol. 67(2):210–220, 1958.
10. D. E. Rutherford. The Cayley-Hamilton theorem for semi-rings. *Proc. Roy. Soc. Edinburgh*, Sect. A 66:211–215, 1964.
11. H. Straubing. A combinatorial proof of the Cayley-Hamilton theorem. *Discrete Maths.*, 43:273–279, 1983.