Determinization of weighted finite automata over strong bimonoids

Miroslav Ćirić, Manfred Droste, Jelena Ignjatović, Heiko Vogler

Faculty of Sciences and Mathematics, University of Niš, Višegradska 33, P. O. Box 224, 18000 Niš, Serbia
Institut für Informatik, Universität Leipzig, D-04009 Leipzig, Germany
Faculty of Computer Science, Technische Universität Dresden, D-01062 Dresden, Germany

Abstract
We consider weighted finite automata over strong bimonoids, where these weight structures can be considered as semirings which might lack distributivity. Then, in general, the well-known run semantics, initial algebra semantics, and transition semantics of an automaton are different. We prove an algebraic characterization for the initial algebra semantics in terms of stable finitely generated submonoids. Moreover, for a given weighted finite automaton we construct the Nerode automaton and Myhill automaton, both being crisp-deterministic, which are equivalent to the original automaton with respect to the initial algebra semantics, resp., the transition semantics. We prove necessary and sufficient conditions under which the Nerode automaton and the Myhill automaton are finite, and we provide efficient algorithms for their construction. Also, for a given weighted finite automaton, we show sufficient conditions under which a given weighted finite automaton can be determinized preserving its run semantics.

Key words: Weighted automaton; Strong bimonoid; Formal power series; Determinization; Nerode automaton; Myhill automaton; Run automaton

1. Introduction

Weighted finite automata are classical nondeterministic automata in which the transitions carry weights. These weights may model, e.g., the amount of resources needed for the execution of a transition, or the probability of its successful execution. The weights often form the algebraic structure of a semiring, and semiring-weighted automata have both a well elaborated theory as well as practical applications, cf. [1, 10, 13, 22, 28, 29, 35].

Recently, a number of authors investigated weighted automata with weights in more general structures, which can be viewed as semirings which might lack distributivity. Examples of such “strong bimonoids” include the real unit interval [0, 1] with t-conorm and t-norm from multivalued logic [20], the “string bimonoid” of all words over an alphabet arising in natural language processing [26], or the algebraic cost structure from algebraic path problems [23]. Important natural examples of strong bimonoids are orthomodular lattices, which serve as a basis of quantum logics [2] where distributivity typically fails; automata based on quantum logics and with weights in orthomodular lattices were investigated in [24, 31, 32, 36, 37, 38]. Automata modeling, e.g., peak power consumption of energy and with particular strong bimonoids as weight structures were recently studied in [7, 8, 9]. Fuzzy automata and fuzzy tree automata defined by a pair of a t-conorm and a t-norm on the real unit interval were investigated in [3, 4] respectively [5], and their study for non-distributive pairs has been appraised as especially interesting.

In this paper, we investigate weighted automata over arbitrary strong bimonoids and, in particular, methods and algorithms for their determinization. This continues the general study of weighted automata...
and weighted tree automata over strong bimonoids begun in [11, 12] respectively [27]. As usual, there are three different definitions of the behavior of such automata: the run semantics [13], the initial algebra semantics [14], and the transition semantics [29]. In the run semantics, the weight of a word \( w \) is computed by summing up the weights of all successful runs of \( \mathcal{A} \) on \( w \) where the weight of a run is the product of the weights of the involved transitions. For the other two semantics, we associate with each input symbol an \( n \times n \)-matrix of transition weights assuming that \( \mathcal{A} \) has \( n \) states. In both semantics, we multiply the vector of initial weights with the matrices corresponding to the symbols of \( w \) and with the final weight vector, but the sequence of these multiplications differs in the two semantics. In the case of semirings, these three semantics coincide for every weighted finite automaton. In contrast to this, we show by an easy example that over the unit interval with bounded sum these three semantics in general differ.

Our main results are the following. We provide algebraic conditions in terms of stable submonoids which are necessary and sufficient for an arbitrary formal power series to be the initial algebra semantics of a weighted finite automaton (cf. Theorem 4.1); this extends a fundamental result from the theory of semiring-weighted automata (cf. [1]). Then we turn to determination results. We are interested in obtaining crisp-deterministic automata: these can be viewed as being deterministic with possibly infinitely many states, but with weights attached only to final states. Given any weighted finite automaton \( \mathcal{A} \), we define three associated automata: the Nerode automaton \( \mathcal{A}_N \), the Myhill automaton \( \mathcal{A}_M \), and, under suitable finiteness conditions on the weights of \( \mathcal{A} \), the run automaton \( \mathcal{A}_r \). Each of them is a crisp-deterministic automaton, and \( \mathcal{A}_r \) is finite. We show that \( \mathcal{A} \) has the same initial algebra, transition, respectively run semantics as \( \mathcal{A}_N \), \( \mathcal{A}_M \), respectively \( \mathcal{A}_r \). We provide necessary and sufficient conditions for \( \mathcal{A}_N \) and for \( \mathcal{A}_M \) to be finite (cf. Theorems 6.3 and 7.4). These conditions depend on the structure of the weights of \( \mathcal{A} \), but do not assume the bimonoid \( K \) to be locally finite. For each case, we develop related algorithms which compute \( \mathcal{A}_N, \mathcal{A}_M \) if they are finite, resp. \( \mathcal{A}_r \). We also prove minimality results for the automata \( \mathcal{A}_N \) and \( \mathcal{A}_M \) in suitable classes of deterministic automata (cf. Theorems 6.6 and 7.8), and we show the structure results that \( \mathcal{A}_N \) is a subdirect product of particular derivative automata, and \( \mathcal{A}_M \) is a subdirect product of particular Nerode automata (cf. Theorems 6.5 and 7.7). For \( \mathcal{A}_N \) and \( \mathcal{A}_M \), these results generalize related results given in [18, 19] for fuzzy automata over complete residuated lattices, and the finiteness result for \( \mathcal{A}_r \) sharpens a result from [11]. We show that the finiteness of \( \mathcal{A}_N \) and \( \mathcal{A}_M \) and the definability of \( \mathcal{A}_r \) are completely independent of each other. Finally, we compute a number of examples demonstrating our results.

The structure of the paper is as follows. In Section 2, we give the concept and many examples of strong bimonoids. In Section 3, we define weighted finite automata over strong bimonoids and their three kinds of semantics. Section 4 contains our algebraic results describing series arising as initial algebra semantics of weighted finite automata. Section 5 contains basic results on crisp-deterministic automata. Section 6 contains our main determination and characterization results for Nerode automata \( \mathcal{A}_N \), Section 7 correspondingly for Myhill automata \( \mathcal{A}_M \), and Section 8 for run automata \( \mathcal{A}_r \). In Section 9 we present simplifications of the results of Section 8 under the assumption that certain elements of the strong bimonoid are additively idempotent. Examples given in Section 10 clarify relationships between finiteness of Nerode and Myhill automata, and definability of run automata, and Section 11 gives the computational examples.

2. Strong bimonoids

Throughout this paper, \( \mathbb{N} \) denotes the set of natural numbers (with zero), \( X^+ \) and \( X^* \) denote respectively the free semigroup and the free monoid over an alphabet \( X \), and \( \varepsilon \) denotes the empty word in \( X^* \).

A bimonoid is a structure \((K, +, \cdot, 0, 1)\) consisting of a set \( K \), two binary operations \( + \) and \( \cdot \) on \( K \) and two constants \( 0, 1 \in K \) such that \((K, +, 0)\) and \((K, \cdot, 1)\) are monoids. As usual, we identify the structure \((K, +, \cdot, 0, 1)\) with its carrier set \( K \). We call \( K \) a strong bimonoid if the operation \( + \) is commutative and \( 0 \) acts as multiplicative zero, i.e., \( a \cdot 0 = 0 = 0 \cdot a \) for every \( a \in K \). We say that a strong bimonoid \( K \) is right distributive, if it satisfies \( (a + b) \cdot c = a \cdot c + b \cdot c \) for every \( a, b, c \in K \); we call \( K \) left distributive, if \( a \cdot (b + c) = a \cdot b + a \cdot c \) for every \( a, b, c \in K \). Then a semiring is a strong bimonoid which is left and right distributive.
Example 2.1. (cf. [11]) (1) The tropical bimonoid is the strong bimonoid \((\mathbb{N}_\infty, +, \min, 0, \infty)\) with \(\mathbb{N}_\infty = \mathbb{N} \cup \{\infty\}\) and the usual extensions of + and min from \(\mathbb{N}\) to \(\mathbb{N}_\infty\). We note that it is not a semiring, because there are \(a, b, c \in \mathbb{N}_\infty\) with \(\min[a+b+c] \neq \min[a,b] + \min[a,c]\) (e.g., take \(a = b = c \neq 0\)).

(2) The tropical semiring is the semiring \((\mathbb{N}_\infty, \min, +, \infty, 0)\).

(3) The algebra \((\{0, 1\}, \oplus, \cdot, 0, 1)\) with the usual multiplication \(\cdot\) of real numbers is a strong bimonoid for, e.g., each of the following two definitions of \(\oplus\) for every \(a, b \in \{0, 1\}\):

\[
    a \oplus b = a + b - a \cdot b \quad \text{(called algebraic sum in [20])}
\]

\[
    a \oplus b = \min(a + b, 1) \quad \text{(called bounded sum in [20]), or: Łukasiewicz t-conorm denoted by \(\mathbb{V}_L\)).}
\]

In neither of the two cases, \((\{0, 1\}, \oplus, \cdot, 0, 1)\) is a semiring.

(4) Let \((C, +, 0)\) be a commutative monoid and let \(A\) be the set of all mappings from \(C\) into itself with pointwise addition, composition of mappings, constant mapping zero, and the identity mapping. Then \(A\) constitutes a strong bimonoid satisfying only one distributivity law (which depends on the order used for defining the composition). Such structures are also called near semirings [21, 34].

(5) Let \(X\) be an alphabet. Consider the strong bimonoid \((X^* \cup \{\infty\}, \wedge, \cdot, \infty, \varepsilon)\) where \(\wedge\) is the longest common prefix operation, \(\cdot\) is the usual concatenation of words, and \(\infty\) is a new element such that \(w \wedge \infty = \infty \wedge w = w\) and \(w \cdot \infty = \infty \cdot w = \infty\) for every \(w \in X^* \cup \{\infty\}\). This bimonoid occurs in investigations for natural language processing, see [26]. It is clear that \((X^* \cup \{\infty\}, \wedge, \cdot, \infty, \varepsilon)\) is left distributive but not right distributive.

(6) The Boolean semiring is the semiring \((\mathbb{B}, \lor, \land, 0, 1)\) with \(\mathbb{B}\) consisting of the truth values 0 and 1, and \(\lor\) and \(\land\) are disjunction and conjunction, respectively.

(7) We note that there are only two strong bimonoids with exactly two elements: the field with two elements and the Boolean semiring (since addition is determined by whether \(1 + 1 = 0\) or \(1 + 1 = 1\)). However, there are strong bimonoids with 3 elements which are not semirings, take, e.g., \((\{0, 1, 2\}, \max, \cdot, 0, 1)\) where \(a \cdot b = (a \cdot b) \mod 3\) for every \(a, b \in \{0, 1, 2\}\).

(8) Bounded lattices (lattices containing a greatest element 1 and a smallest element 0) are strong bimonoids. As is well known, there are large classes of lattices that are not distributive [15].

(9) Moreover, bounded distributive lattices, semiring-reducts of semilattice ordered monoids and of complete residuated lattices, and Brouwerian lattices are semirings.

Recall that an algebra is locally finite if any of its finitely generated subalgebras is finite. The strong bimonoid \((K, +, \cdot, 0, 1)\) is additively locally finite if the monoid \((K, +, 0)\) is locally finite, and it is multiplicatively locally finite if the monoid \((K, \cdot, 1)\) is locally finite. If \(K\) is both additively and multiplicatively locally finite, then it is called bi-locally finite.

Example 2.2. The algebra \((\{0\} \cup \{\lambda, 1\}, +, \cdot, 0, 1)\), for \(\lambda < \frac{1}{2}\), with the product of \(a\) and \(b\) being the usual one \(a \cdot b\) if \(a \cdot b \geq \lambda\), and 0 if \(a \cdot b < \lambda\), and addition \(a + b = \min\{a + b, 1\}\), is a bi-locally finite strong bimonoid, but not a locally finite strong bimonoid.

Let \(P\) and \(Q\) be sets. We let \(Q^P\) denote the set of all functions from \(P\) to \(Q\). Next, let \(K\) be a strong bimonoid and let \(A\) be a finite non-empty set. A mapping \(\mu : A \times A \rightarrow K\) is called an \(A\times A\)-matrix over \(K\), and a mapping \(\nu : A \rightarrow K\) is called an \(A\)-vector over \(K\). If \(K\) is a particular ordered set (e.g., the real unit interval \([0, 1]\)), then matrices are called fuzzy relations, and vectors are called fuzzy subsets in the literature.

Given matrices \(\mu_1, \mu_2 \in K^{A \times A}\) and vectors \(v_1, v_2 \in K^A\). Then we define the matrix product \(\mu_1 \cdot \mu_2 \in K^{A \times A}\), the matrix-vector products \(v_1 \cdot \mu_1 \in K^A\) and \(\mu_1 \cdot v_1 \in K^A\), and the scalar product \(v_1 \cdot v_2 \in K\) as follows for every \(a_1, a_2 \in A\):

\[
    (\mu_1 \cdot \mu_2)(a_1, a_2) = \sum_{a \in A} \mu_1(a_1, a) \cdot \mu_2(a, a_2),
\]

\[
    (v_1 \cdot \mu_1)(a_1) = \sum_{a \in A} v_1(a) \cdot \mu_1(a, a_1),
\]

\[
    (\mu_1 \cdot v_1)(a_1) = \sum_{a \in A} \mu_1(a_1, a) \cdot v_1(a),
\]

\[
    v_1 \cdot v_2 = \sum_{a \in A} v_1(a) \cdot v_2(a).
\]
Recall that the addition of \( K \) is commutative and that \( A \) is non-empty; thus, the sums on the right-hand sides are well defined. Since strong bimonoids lack distributivity of the multiplication operation over the addition operation, the matrix product and matrix-vector products need not be associative. However, if \( K \) is a semiring, then they are associative.

Let \((S, \cdot)\) be a semigroup. The least subsemigroup of \((S, \cdot)\) which contains an element \( a \in S \) is called the monogenic subsemigroup of \((S, \cdot)\) generated by \( a \), and denoted by \( \langle a \rangle \). If \( \langle a \rangle \) is finite, then there exists the least number \( i \in \mathbb{N}\backslash\{0\} \) such that \( a^i = a^{i+k} \), for some \( k \in \mathbb{N}\backslash\{0\} \), and there exists the least number \( p \in \mathbb{N}\backslash\{0\} \) such that \( a^p = a^{i+p} \). We call \( i = i(a) \) the index of \( a \), and \( p = p(a) \) the period of \( a \). In this case the number of elements of \( \langle a \rangle \) is \( i(a) + p(a) - 1 \), this number is called the order of \( a \), and we say that \( a \) has a finite order. Otherwise, if \( \langle a \rangle \) is infinite, we say that \( a \) has an infinite order. It is well-known that in this case the semigroup \((\langle a \rangle, \cdot)\) is isomorphic to \((\mathbb{N}\backslash\{0\},+)\).

### 3. Weighted finite automata and series over strong bimonoids

In the rest of this paper, if not noted otherwise, let \( K \) be a strong bimonoid and \( X \) an alphabet.

A formal power series over \( X \) and \( K \), for short a series, is any mapping \( \varphi: X^* \to K \). Instead of \( \varphi(u) \) we write \( \varphi \cdot u \) for every \( u \in X^* \). The set of all series over \( X \) and \( K \) is denoted by \( K\langle X^* \rangle \). The image of \( \varphi \) is the set \( \text{Im}(\varphi) = \{ \varphi \cdot u \mid u \in X^* \} \).

A weighted finite automaton (short: wfa) over \( X \) and \( K \) is a quadruple \( \mathcal{A} = (A, \delta, \sigma, \tau) \), where \( A \) is a finite non-empty set of states, \( \delta: A \times X \times A \to K \) is a weighted transition function, \( \sigma: A \to K \) is an initial weight vector and \( \tau: A \to K \) is a final weight vector. For each \( x \in X \) we define a weighted transition matrix (or a weighted transition relation) \( \delta_x: A \times A \to K \) by \( \delta_x(a,b) = \delta(a,x,b) \), for all \( a,b \in A \). If \( \mathcal{A} \) is a wfa, we will write \( |\mathcal{A}| = \left| A \right| \), i.e., \( |\mathcal{A}| \) denotes the number of states of \( \mathcal{A} \).

Due to lack of distributivity in \( K \), we define the behavior of a wfa in several different ways. In particular, we distinguish the initial algebra semantics, the run semantics, and the transition semantics of a weighted finite automaton [11].

**Initial algebra semantics:** For each \( u \in X^* \) we define a vector \( \sigma_u: A \to K \) inductively, as follows: \( \sigma_a = \sigma \), and for all \( u \in X^* \) and \( x \in X \) we set \( \sigma_{ux} = \sigma_u \cdot \delta_x \). Clearly, if \( u = x_1x_2 \cdots x_n \), where \( x_1, \ldots, x_n \in X \), then

\[
\sigma_u = (\ldots((\sigma \cdot \delta_{x_1}) \cdot \delta_{x_2}) \cdots) \cdot \delta_{x_n}.
\]

The \( i \)-behavior of \( \mathcal{A} \) is the series \( [\mathcal{A}]_i \) in \( K\langle X^* \rangle \) defined by

\[
([\mathcal{A}]_i) \cdot u = \sigma_u \cdot \tau = \left( (\ldots((\sigma \cdot \delta_{x_1}) \cdot \delta_{x_2}) \cdots) \cdot \delta_{x_n} \right) \cdot \tau.
\]

for every \( u = x_1x_2 \cdots x_n \in X^* \), where \( x_1, x_2, \ldots, x_n \in X \).

**Transition semantics:** For each \( u \in X^* \) we define a matrix \( \delta_u: A \times A \to K \) inductively, as follows: \( \delta_e \) is the unit matrix, and for all \( u \in X^* \) and \( x \in X \) we set \( \delta_{ux} = \delta_u \cdot \delta_x \). In other words, if \( u = x_1x_2 \cdots x_n \in X^* \), where \( x_1, x_2, \ldots, x_n \in X \), then

\[
\delta_u = (\ldots((\delta_{x_1} \cdot \delta_{x_2}) \cdot \delta_{x_3}) \cdots) \cdot \delta_{x_n}.
\]

Matrices \( \delta_u, u \in X^* \), will be also called weighted transition matrices.

The \( t \)-behavior of \( \mathcal{A} \) is the series \( [\mathcal{A}]_t \) in \( K\langle X^* \rangle \) defined by

\[
([\mathcal{A}]_t) \cdot u = (\sigma \cdot \delta_u) \cdot \tau.
\]

for every \( u \in X^* \). If \( K \) is a semiring, then the matrix product and matrix-vector product (whenever defined) are associative and the above parentheses can be deleted. In this case, the \( t \)-behavior becomes the \( f \)-behavior defined in [11].

**Run semantics:** The \( r \)-behavior of \( \mathcal{A} \) is the series \( [\mathcal{A}]_r \) in \( K\langle X^* \rangle \) defined by

\[
([\mathcal{A}]_r) \cdot u = \sum_{(a_0,a_1,\ldots,a_n) \in A^{n+1}} \sigma(a_0) \cdot \delta_{x_1}(a_0,a_1) \cdot \delta_{x_2}(a_1,a_2) \cdots \delta_{x_n}(a_{n-1},a_n) \cdot \tau(a_n).
\]
for every \( u = x_1 x_2 \cdots x_n \in X^* \), where \( x_1, x_2, \ldots, x_n \in X \). Here, a sequence \( (a_0, a_1, \ldots, a_n) \in A^{n+1} \) is often called a \textit{run} of \( A \) on \( u = x_1 x_2 \cdots x_n \).

Let \( s \in \{i, t, r\} \). A series \( \varphi \in \mathcal{K}(X^*) \) is called \textit{s-recognizable} if there exists a wfa \( A \) over \( X \) and \( K \) such that \( [A]_s = \varphi \). We say that two wfa \( A \) and \( A' \) over \( X \) and \( K \) are \textit{s-equivalent} if \( [A]_s = [A']_s \).

Example 3.1. Let \( K = ([0, 1], \oplus, \cdot, 0, 1) \) be the strong bimonoid given in Example 2.1 (3), where \( \oplus \) is the bounded sum, and let \( A \) be a wfa over an alphabet \( X = \{x, y\} \) and \( K \) given by the graph in Figure 1 where, as usual, transitions with weights 0 or 1 are not shown. Then

\[
([A]_s, xy) = 0.6, \quad ([A]_i, xy) = 0.36, \quad ([A]_r, xy) = 0.72,
\]

and therefore, all three semantics of the wfa \( A \) are different.

Let \( A = (A, \delta, \sigma, \tau) \) be a wfa over \( X \) and \( K \). For any \( \sigma' : A \to K \), the wfa \( (A, \delta, \sigma', \tau) \) will be called an \textit{initial variant} of \( A \). Likewise, for any \( \tau' : A \to K \), the wfa \( (A, \delta, \sigma, \tau') \) will be called a \textit{final variant} of \( A \), and for any \( \sigma', \tau' : A \to K \), the wfa \( (A, \delta, \sigma', \tau') \) will be called a \textit{variant} of \( A \). By \( \text{Rec}(A, \sigma) \) we denote the family of all series in \( \mathcal{K}(X^*) \) which are \textit{i-behaviors} of some final variant of \( A \), and by \( \text{Rec}(A) \) we denote the family of all series in \( \mathcal{K}(X^*) \) which are \textit{t-behaviors} of some variant of \( A \).

4. Necessary and sufficient conditions for \textit{i-recognizability}

In the theory of weighted automata over a semiring \( K \) there is a fundamental algebraic characterization of when a series \( \varphi \in \mathcal{K}(X^*) \) is recognizable: this is the case if and only if \( \varphi \) belongs to some finitely generated stable subsemimodule of \( \mathcal{K}(X^*) \) (cf. [1]). In this section we wish to investigate such implications for strong bimonoids and \( i \)-recognizability.

Let \( K \) be a strong bimonoid. We call a commutative monoid \( (M, +, 0) \) a \textit{K-monoid}, if \( K \) operates on \( M \) (from the right), i.e., there is a multiplication \( \cdot : M \times K \to K \) such that for all \( m \in M \) and \( k, k' \in K \) : \( m \cdot (k k') = (m \cdot k) \cdot k' \), \( 0 \cdot k = 0, m \cdot 0 = 0 \) and \( m \cdot 1 = m \). We say that this operation is

- right distributive, if \( (m + m') \cdot k = m \cdot k + m' \cdot k \), for all \( m, m' \in M \) and \( k \in K \);
- left distributive, if \( m \cdot (k + k') = m \cdot k + m \cdot k' \), for all \( m \in M \) and \( k, k' \in K \).

Now, let \( M \) be a \( K \)-monoid. Let \( S \) be a submonoid of \( M \) and let \( F \) be a subset of \( S \). We say that \( S \) is \( K \)-\textit{generated} by \( F \), if \( S \) is the submonoid of \( M \) generated by the set \( \{ f \cdot k \mid f \in F, k \in K \} \), that is, each \( s \in S \) can be written in the form \( s = \sum_{i=1}^{m} f_i \cdot k_i \) for some \( m \geq 0, f_1, \ldots, f_m \in F \), and \( k_1, \ldots, k_m \in K \). Clearly, if in this case the operation of \( K \) on \( M \) is right distributive, then \( S \) is a \( K \)-monoid, i.e., \( s \cdot k \in S \) for all \( s \in S \) and \( k \in K \). We say that \( S \) is \( K \)-\textit{finitely generated}, if \( S \) is \( K \)-generated by some finite subset \( F \subseteq S \).

Now assume that \( S \) is \( K \)-generated by the finite set \( F = \{ f_1, \ldots, f_n \} \) and that the operation of \( K \) on \( M \) is left distributive. We claim that then we can write each \( s \in S \) in the form

\[
s = \sum_{i=1}^{n} f_i \cdot k_i \quad \text{with} \quad k_1, \ldots, k_n \in K.
\]
Indeed, write
\[ s = \sum_{j=1}^{m} f_j \cdot k_j \] with \( m \geq n \), \( f_1, \ldots, f_m \in F \) and \( k_1, \ldots, k_m \in K \).

We may assume that \( m \geq n \), and that each \( f_i \in F \) occurs as some \( f'_i \) (possibly, with \( k'_i = 0 \)). Now if \( m > n \) and e.g. \( f'_1 = f'_2 \), observe that \( f'_1 \cdot k'_1 + f'_2 \cdot k'_2 = f'_1(k'_1 + k'_2) \) by left distributivity. Hence we may simplify this sum to obtain the description of \( s \) as in (6).

Now let \( X \) be an alphabet. Note that \( (K \langle X \rangle, +, 0) \) is a commutative monoid, and it becomes a \( K \)-monoid by the usual pointwise multiplication from the right; that is, for \( \varphi \in K \langle X \rangle, k \in K \), and \( u \in X^* \) let \( (\varphi \cdot k, u) = (\varphi, u) \cdot k \). We note that if the bimonoid \( K \) is left (resp. right) distributive, then this operation of \( K \) on \( K \langle X \rangle \) is also left (resp. right) distributive.

Given \( \varphi \in K \langle X \rangle \) and \( v \in X^* \), we define the series \( v^{-1} \varphi \in K \langle X \rangle \) and \( \varphi v^{-1} \in K \langle X \rangle \) by letting \( (v^{-1} \varphi, u) = (\varphi, vu) \) and \( (\varphi v^{-1}, u) = (\varphi, uv) \) for \( u \in X^* \). The series \( v^{-1} \varphi \) is called a \textit{left derivative}, and the series \( \varphi v^{-1} \) a \textit{right derivative} of \( \varphi \) with respect to \( v \). Derivatives of series have been well investigated for series over semirings (cf. [1, Section I.5]).

The following rules are easy to check for all \( \varphi, \varphi' \in K \langle X \rangle \), \( u, v \in X^* \) and \( k \in K \):

\[
\begin{align*}
(\varphi + \varphi')v^{-1} &= \varphi v^{-1} + \varphi' v^{-1}, & (1) \\
(\varphi \cdot k)v^{-1} &= (\varphi v^{-1}) \cdot k, & (2) \\
\varphi (uv)^{-1} &= (\varphi v^{-1})u^{-1}. & (3)
\end{align*}
\]

Analogous rules hold for left derivatives. We say that a subset \( S \subseteq K \langle X \rangle \) is \textit{stable} if \( \varphi v^{-1} \in S \) for all \( \varphi \in S \) and \( v \in X^* \).

Now we can show:

**Theorem 4.1.** Consider the following two conditions for \( \varphi \in K \langle X \rangle \):

1. \( \varphi \) is \( i \)-recognizable.
2. \( \varphi \) is contained in some stable \( K \)-finitely generated submonoid of \( K \langle X \rangle \).

(a) If \( K \) is right distributive, then (1) \( \Rightarrow \) (2).
(b) If \( K \) is left distributive, then (2) \( \Rightarrow \) (1).

**Proof.** (a) Let \( \varphi = [\mathcal{A}] \), for some wfa \( \mathcal{A} = (A, \delta, \sigma, \tau) \) over \( X \) and \( K \). For each \( a \in A \) we define the series \( \psi_a \in K \langle X \rangle \) by letting \( (\psi_a, u) = \sigma_a(a) \) for each \( u \in X^* \). Let \( S \) be the submonoid of \( K \langle X \rangle \) which is \( K \)-finitely generated by \( \{ \psi_a \mid a \in A \} \). We show that \( S \) is stable.

Indeed, let \( a \in A \) and \( x \in X \). We first show that \( \psi_a x^{-1} \in S \). For any \( u \in X^* \) we have

\[
(\psi_a x^{-1}, u) = (\psi_a, ux) = \sigma_{ax}(a) = \sum_{b \in A} \sigma_a(b) \cdot \delta_x(b, a) = \sum_{b \in A} (\psi_b, u) \cdot \delta_x(b, a),
\]

so \( \psi_a x^{-1} = \sum_{b \in A} \psi_b \cdot \delta_x(b, a) \in S \). Since \( K \) is right distributive, the operation of \( K \) on \( K \langle X \rangle \) is right-distributive and \( S \) is a \( K \)-monoid. Together with rules (7) and (8), this shows that \( \psi x^{-1} \in S \), for every \( \psi \in S \) and \( x \in X \). Now rule (9) implies that \( S \) is stable.

It remains to show that \( \varphi \in S \). For any \( u \in X^* \) we have

\[
(\varphi, u) = ([\mathcal{A}], u) = \sigma_{\mathcal{A}} \cdot \tau = \sum_{a \in A} \sigma_a(a) \cdot \tau_{\mathcal{A}} = \sum_{a \in A} (\psi_a, u) \cdot \tau(a),
\]

so \( \varphi = \sum_{a \in A} \psi_a \cdot \tau(a) \in S \).

(b) Let \( S \) be a stable \( K \)-finitely generated submonoid of \( K \langle X \rangle \) such that \( \varphi \in S \). Choose a finite \( K \)-generating subset \( F = \{ \varphi_1, \ldots, \varphi_m \} \) of \( S \). We define a wfa \( \mathcal{A} = (A, \delta, \sigma, \tau) \) as follows. We put \( A = \{ 1, \ldots, n \} \).
Since the operation of $K$ on $K\langle X'\rangle$ is left distributive, we can write $\varphi \in S$ in the form $\varphi = \sum_{i=1}^{n} \varphi_{i} \cdot k_{i}$ with $k_{1}, \ldots, k_{n} \in K$. We define $\sigma$ and $\tau$ by letting $\sigma(i) = (\varphi_{i}, v)$ and $\tau(i) = k_{i}$ for each $i \in A$.

To define $\delta$, choose $x \in X$ and $j \in A$. Since $S$ is stable, we have $\varphi_{j} x^{-1} \in S$, so $\varphi_{j} x^{-1} = \sum_{i=1}^{n} \varphi_{i} \cdot k'_{i}$ for some $k_{1}', \ldots, k'_{n} \in K$ (depending on $x$ and $j$). Then we put $\delta(x, j) = k'_{i}$ for each $i \in A$.

We claim that $[\mathcal{A}]_{\varphi}$. First we show that $\sigma(i) = (\varphi_{i}, u)$ for each $i \in A$ and $u \in X$. We proceed by induction on the length of $u$. Clearly, $\sigma(i) = \sigma(0) = (\varphi_{0}, v)$. Now, let $u \in X$ such that $\sigma(i) = (\varphi_{i}, u) = (\varphi_{i} u^{-1}, v)$ for each $i \in A$, and let $x \in X$. Then for each $j \in A$ by rules (7)–(9) we have

$$
\sigma_{w}(j) = \sum_{i \in A} \sigma_{i}(i) \cdot \delta_{x}(i, j) = \sum_{i \in A} (\varphi_{i} u^{-1}, v) \cdot \delta_{x}(i, j) = \left( \sum_{i \in A} \varphi_{i} \cdot \delta_{x}(i, j) \right) u^{-1}, v
$$

as needed. Now let $u \in X$. Then

$$
([\mathcal{A}]_{\varphi}, u) = \sigma_{u} \cdot \tau = \sum_{i \in A} \sigma_{i}(i) \cdot \tau(i) = \sum_{i \in A} (\varphi_{i}, u) \cdot k_{i} = (\varphi, u),
$$

proving our claim. □

Let $\varphi \in K\langle X'\rangle$. The Hankel matrix $H(\varphi)$ of $\varphi$ is defined to be the infinite $X' \times X'$-matrix with entries from $K$ given by $H(\varphi)(u, v) = (\varphi, v)$ for $u, v \in X'$. Note that the $v$-column $H(\varphi)(-, v)$ of $H(\varphi)$ equals the series $\varphi v^{-1}$. The next result generalizes [29, Cor. II.3.2].

**Corollary 4.2.** Let $K$ be a finite strong bimoid and $\varphi \in K\langle X'\rangle$. Consider the following two conditions.

(1) $\varphi$ is $i$-recognizable.
(2) $H(\varphi)$ has only finitely many pairwise different columns.

(a) If $K$ is right distributive, then (1) $\Rightarrow$ (2).
(b) If $K$ is left distributive, then (2) $\Rightarrow$ (1).

**Proof.** (a) We follow the proof of Theorem 4.1 (a).

The submonoid $S$ of $K\langle X'\rangle$ is generated by the finite set $\{\varphi_{a} \cdot k \mid a \in A, k \in K\}$. Since $K$ is finite, each finitely generated submonoid of $(K\langle X'\rangle, +, 0)$ is finite. By the proof of Theorem 4.1, $S$ is stable and $\varphi \in S$, so $\varphi v^{-1} \in S$ for all $v \in X'$. Hence $S$ contains all columns of $H(\varphi)$ which therefore constitute a finite set.

(b) Let $S$ be the submonoid of $K\langle X'\rangle$ which is $K$-generated by the columns of $H(\varphi)$, i.e., by all series $\varphi v^{-1}$ where $v \in X'$. By rules (7)–(9), clearly $S$ is stable. Now Theorem 4.1 (b) implies that $\varphi$ is $i$-recognizable. □

5. **Crisp-deterministic weighted finite automata**

Let $\mathcal{A} = (A, \delta, \sigma, \tau)$ be a wfa over $X$ and $K$. The weighted transition function $\delta$ is called **crisp-deterministic** if for every $x \in X$ and every $a \in A$ there exists $a' \in A$ such that $\delta_{x}(a, a') = 1$, and $\delta_{x}(a, b) = 0$, for all $b \in A \setminus \{a'\}$. The initial weight vector $\sigma$ is called **crisp-deterministic** if there exists $a_{0} \in A$ such that $\sigma(a_{0}) = 1$ and $\sigma(a) = 0$ for every $a \in A \setminus \{a_{0}\}$. If both $\sigma$ and $\delta$ are crisp-deterministic, then $\mathcal{A}$ is called a **crisp-deterministic weighted finite automaton** (for short: **cdfa**).

Equivalently, we can define a crisp-deterministic weighted finite automaton over $X$ and $K$ as a quadruple $\mathcal{A} = (A, \delta, a_{0}, \tau)$, where $A$ is a non-empty set of states, $\delta : A \times X \rightarrow A$ is a transition function, $a_{0} \in A$ is an initial state and $\tau : A \rightarrow K$ is a final weight vector. The transition function $\delta$ can be extended to a function $\delta' : A \times X' \rightarrow A$ as follows: $\delta'(a, x) = a$, for each $a \in A$, and $\delta'(a, ux) = \delta(\delta'(a, u), x)$, for every $a \in A$, $u \in X'$ and $x \in X$. Also, we allow the set $A$ to be infinite, and then $\mathcal{A}$ is called a **crisp-deterministic weighted automaton** (for short: **cdfa**). The **behavior** of $\mathcal{A}$ is the series $[\mathcal{A}]$ in $K\langle X'\rangle$ defined by

$$
([\mathcal{A}], u) = \tau(\delta'(a_{0}, u)).
$$

(10)
for every \( u \in X^* \). Obviously, the image of \( \|A\| \) is contained in the image of \( \tau \) which is finite if the set of states \( A \) is finite. By \( |A| \) we denote the cardinality of the set of states of \( A \), i.e., \( |A| = |A| \). We extend the concept of \( s \)-equivalent wfa in the obvious way to include also cdwa.

A series \( \phi \in K(\mathcal{X}^*) \) is called cdwfa-recognizable if there exists a crisp-deterministic weighted finite automaton \( A \) over \( X \) and \( K \) such that \( \|A\| = \phi \). We also say that \( A \) recognizes \( \phi \).

Let us note that the above concept of a cdwfa slightly differs from the concept having the same name defined in [11]. Namely, the weighted transition function \( \delta \) was defined in [11] to be crisp-deterministic if for every \( x \in X \) and \( a \in A \) either there exists \( a' \in A \) such that \( \delta_x(a, a') = 1 \) and \( \delta_x(b) = 0 \) for all \( b \in A \setminus \{a'\} \), or \( \delta_x(a, b) = 0 \), for all \( b \in A \). In other words, \( \delta \) was defined in [11] to be crisp-deterministic if it can be regarded as a partial function from \( A \times X \) to \( K \), and here we require that this function is total. Clearly, this is no essential difference, as any partial function \( \delta \) can be completed by adding a dummy state to \( X \). Also, the initial weight vector \( \sigma \) was defined to be crisp-deterministic if either there exists \( a_0 \in A \) such that \( \sigma(a_0) = 1 \) and \( \sigma(a) = 0 \) for every \( a \in A \setminus \{a_0\} \), or \( \sigma(a) = 0 \), for all \( a \in A \).

**Theorem 5.1.** Let \( A = (A, \delta, \sigma, \tau) \) be a weighted finite automaton.

(a) If either \( \sigma \) or \( \delta \) is crisp-deterministic, then for every \( u \in X^* \) we have that \( \sigma_u = \sigma \cdot \delta_u \), and consequently, \( \|A\| = \|A\| \).

(b) If \( A \) is crisp-deterministic, then \( \|A\| = \|A\| \cdot \|A\| = \|A\| \).

**Proof.** Without loss of generality we can assume that \( A = \{1, 2, \ldots, n\} \).

(a) Assume first that \( \sigma \) is crisp-deterministic. Then there exists \( i_0 \in A \) such that \( \sigma(i_0) = 1 \) and \( \sigma(i) = 0 \) for each \( i \in A \setminus \{i_0\} \). We will prove the equality \( \sigma_u = \sigma \cdot \delta_u \) by induction on the length of \( u \). Clearly, for the empty word we have that \( \sigma \cdot \delta_e = \sigma = \sigma_e \). Suppose that \( \sigma_u = \sigma \cdot \delta_u \), for some \( u \in X^* \), and consider arbitrary \( x \in X \) and \( k \in A \). Then

\[
[\sigma \cdot \delta_u \cdot \delta_x](k) = \sum_{j=1}^{n} [\sigma \cdot \delta_u](j) \cdot \delta_x(j, k) = \sum_{j=1}^{n} \sum_{i=1}^{n} \sigma(i) \cdot \delta_u(i, j) \cdot \delta_x(j, k) = \delta_x(i_0, k) = \sum_{i=1}^{n} \sigma(i) \cdot \delta_u(i, k) = (\sigma \cdot \delta_u)(k),
\]

so \( \sigma \cdot \delta_u \cdot \delta_x = \sigma \cdot \delta_u \), and therefore, \( \sigma_u \cdot \delta_x = \sigma \cdot \delta_u \cdot \delta_x = (\sigma \cdot \delta_u) \cdot \delta_x = \sigma \cdot \delta_u \cdot \delta_x \). Hence, by induction we obtain that \( \sigma_u \cdot \delta_x \) holds for every \( u \in X^* \).

Assume now that \( \delta \) is crisp-deterministic. For arbitrary \( k \in A \) and \( u \in X^* \) set \( I_u^k = \{ j \in A \mid \delta_u(j, k) = 1 \} \). Since \( \delta \) is a crisp-deterministic function, we have that \( I_u^k \cap I_v^k = \emptyset \), for all \( u \in X^* \) and \( k, l \in A, k \neq l \). Next we prove that

\[
I_u^k = \bigcup_{j \in I_u^k} I_j^l,
\]

for all \( k \in A \), \( u \in X^* \) and \( x \in X \). Indeed, if \( i \in I_u^k \), then there exists \( j \in A \) such that \( \delta_u(i, j) = 1 \) and \( \delta_x(j, k) = 1 \), which means that \( j \in I_u^k \) and \( i \in I_j^l \). Hence, \( i \in \bigcup_{j \in I_u^k} I_j^l \). Conversely, let \( i \in I_j^l \), for some \( j \in I_u^k \). Then \( \delta_u(i, j) = 1 \) and \( \delta_x(j, k) = 1 \), and since \( \delta \) is a crisp-deterministic function, we obtain that \( \delta_u(i, k) = 1 \), so \( i \in I_u^k \). Thus, (11) holds.

Now we prove the equality \( \sigma_u \cdot \delta_x \cdot \delta_u \) by induction on the length of \( u \). Clearly, for the empty word we have that \( \sigma \cdot \delta_e = \sigma = \sigma_e \). Suppose that \( \sigma_u \cdot \delta_x \cdot \delta_u \), for some \( u \in X^* \), and consider arbitrary \( x \in X \) and \( k \in A \). Then (11) we obtain that

\[
[\sigma \cdot \delta_u \cdot \delta_x](k) = \sum_{j=1}^{n} [\sigma \cdot \delta_u](j) \cdot \delta_x(j, k) = \sum_{j=1}^{n} \sum_{i=1}^{n} \sigma(i) \cdot \delta_u(i, j) \cdot \delta_x(j, k) = \sum_{i=1}^{n} \sigma(i) \cdot \delta_u(i, k) = (\sigma \cdot \delta_u)(k),
\]

...
so \((\sigma \cdot \delta_u) \cdot \delta_x = \sigma \cdot \delta_{ux}\). Now, by the induction hypothesis we obtain that \(\sigma_{ux} = \sigma \cdot \delta_x = (\sigma \cdot \delta_u) \cdot \delta_x = \sigma \cdot \delta_{ux}\). Thus, by induction we conclude that \(\sigma_u = \sigma \cdot \delta_u\) holds for every word \(u \in X\).

(b) By (a) we obtain that \([A] = \prod [A], = \prod [A], \) and we have \([A] = \prod [A], \) by Remark 7 of [11].

At the end of this section we define some notions concerning cdwa, which will be used in the rest of the paper.

Let \(\mathcal{A} = (A, \delta, a_0, \tau)\) be a cdwa. If \(A'\) is a subset of \(A\) such that \(a_0 \in A'\) and \(\delta(a, x) \in A', \) for all \(a \in A'\) and \(x \in X, \) and if \(\delta' : A' \times X \to A'\) is the restriction of \(\delta\) to \(A' \times X, \) \(a_0' = a_0, \) and \(\tau' : A' \to K\) is the restriction of \(\tau\) to \(A', \) then \(\mathcal{A}' = (A', \delta', a_0', \tau')\) is also a cdwa, and it is called a subautomaton of \(\mathcal{A}.\) A state \(a \in A\) is called accessible if there exists \(u \in X^*\) such that \(\delta'(a_0, u) = a.\) If every state of \(\mathcal{A}\) is accessible, then \(\mathcal{A}\) is called an accessible cdwa. Also, if \(A'\) is the set of all accessible states of \(\mathcal{A},\) and if \(\delta' : A' \times X \to A'\) is the restriction of \(\delta\) to \(A' \times X, a_0' = a_0,\) and \(\tau' : A' \to K\) is the restriction of \(\tau\) to \(A',\) then the subautomaton \(\mathcal{A}' = (A', \delta', a_0', \tau')\) of \(\mathcal{A}\) is an accessible cdwa, called the accessible part of \(\mathcal{A}.\)

Let \(\mathcal{A} = (A, \delta, a_0, \tau)\) and \(\mathcal{A}' = (A', \delta', a_0', \tau')\) be cdwas. If \(\phi : A \to A'\) is a bijective mapping satisfying \(\phi(a_0) = a_0', \phi(\delta(a, x)) = \delta'(\phi(a), x),\) for all \(a \in A\) and \(x \in X,\) and \(\tau(a) = \tau'(\phi(a)),\) for each \(a \in A,\) then \(\phi\) is called an isomorphism of these cdwas, and we say that \(\mathcal{A}\) and \(\mathcal{A}'\) are isomorphic cdwas.

Let \(\mathcal{A} = (A, \delta, a_0, \tau)\) be a cdwa. An equivalence relation \(\mu\) on \(A\) is a congruence on \(\mathcal{A}\) if \((a, b) \in \mu\) implies \(\tau(a) = \tau(b)\) and \(\delta(a, x) = \delta(b, x)\) \(\in \mu,\) for all \(a, b \in A\) and \(x \in X\) (cf. [19]). If \(\mu\) is a congruence on \(\mathcal{A}\) and \(A/\mu = A/\mu\) is the related factor set, then we define mappings \(\delta_\mu : A_\mu \times X \to A_\mu\) and \(\tau_\mu : A_\mu \to K\) by \(\delta_\mu([a]_\mu) = [\delta(a, x)]_\mu\) and \(\tau_\mu([a]_\mu) = \tau(a),\) for all \(a \in A\) and \(x \in X,\) and the quadruple \(\mathcal{A}_\mu = (A_\mu, \delta_\mu, [a_0]_\mu, \tau_\mu)\) is a cdwa, called the factor automaton of \(\mathcal{A}\) with respect to \(\mu.\)

Let \(\mathcal{A}_i = (A_i, \delta_i, a_{0,i}, \tau_i),\) for \(i \in \{1, \ldots, n\},\) be a collection of cdwas, let \(A_{11} = A_1 \times \cdots \times A_n,\) let \(a_0 = (a_{0,1}, \ldots, a_{0,n}),\) and let \(\delta_{11} : A_{11} \times X \to A_{11}\) and \(\tau_{11} : A_{11} \to K\) be mappings defined by

\[
\delta_{11}(a_1, \ldots, a_n, x) = (\delta_1(a_1, x), \ldots, \delta_n(a_n, x)),
\]

\[
\tau_{11}(a_1, \ldots, a_n) = \tau_1(a_1) \cdot \cdots \cdot \tau_n(a_n),
\]

for all \((a_1, \ldots, a_n) \in A_{11}\) and \(x \in X.\) Then \(\mathcal{A}_{11} = (A_{11}, \delta_{11}, a_{0,11}, \tau_{11})\) is also a cdwa, and it is called the direct product of cdwas \(\mathcal{A}_i, i \in \{1, \ldots, n\}.\) For any subset \(A'_{11}\) of \(A_{11}\) and any \(i \in \{1, \ldots, n\},\) the mapping \(p_i : A'_{11} \to A_i\) defined by \(p_i(a_1, \ldots, a_n) = a_i,\) for each \((a_1, \ldots, a_n) \in A'_{11}\) is called the projection mapping of \(A'_{11}\) into \(A_i.\) Any subautomaton \(\mathcal{A}'_{11} = (A'_{11}, \delta'_{11}, a'_{0,11}, \tau'_{11})\) of \(\mathcal{A}_{11}\) having the property that for each \(i \in \{1, \ldots, n\}\) the projection mapping \(p_i : A'_{11} \to A_i\) is surjective, is called a subdirect product of cdwa \(\mathcal{A}_i, i \in \{1, \ldots, n\}.\) Observe that the accessible part of \(\mathcal{A}_{11}\) is a subdirect product of the cdwa \(\mathcal{A}_i (i \in \{1, \ldots, n\})\) if and only if each \(\mathcal{A}_i (i \in \{1, \ldots, n\})\) is accessible. It is worth noting that direct and subdirect products of cdwas have a natural interpretation as parallel compositions of these automata.

For a series \(\varphi \in K(X^*)_1\) let \(A_\varphi = \{u^{-1}\varphi \mid u \in X^*\}\) denote the set of all left derivatives of \(\varphi,\) and let \(\delta_\varphi : A_\varphi \times X \to A_\varphi\) and \(\tau_\varphi : A_\varphi \to K\) be mappings defined by

\[
\delta_\varphi(\psi, x) = x^{-1}\psi \quad \text{and} \quad \tau_\varphi(\psi) = (\psi, \epsilon),
\]

for every \(\psi \in A_\varphi\) and \(x \in X.\) Then \(\mathcal{A}_\varphi = (A_\varphi, \delta_\varphi, \varphi, \tau_\varphi)\) is an accessible cdwa, and it is called the derivative automaton of the series \(\varphi\) [17, 19]. It has been proved in [19] that the derivative automaton \(\mathcal{A}_\varphi\) is finite if and only if the series \(\varphi\) is cdwa-recognizable. Namely, \(\mathcal{A}_\varphi\) is a minimal cdwa which recognizes \(\varphi\) [19]. An algorithm for construction of the derivative automaton of a series, based on simultaneous construction of the derivative automata of ordinary languages \(\varphi^{-1}(a),\) for all \(a \in \operatorname{Im}(\varphi),\) has been also given in [19].

6. Determinization related to the initial algebra semantics

Let \(\mathcal{A} = (A, \delta, a, \tau)\) be a wfa over \(X\) and \(K.\) Let us set \(A_N = \{a_u \mid u \in X^*\},\) and let us define \(\delta_N : A_N \times X \to A_N\) and \(\tau_N : A_N \to K\) by

\[
\delta_N(a_u, x) = a_{ux} \quad \text{and} \quad \tau_N(a_u) = a_u \cdot \tau,
\]

for every \(u \in X^*\) and \(x \in X.\)
Proposition 6.1. Let \( \mathcal{A} = (A, \delta, \sigma, \tau) \) be a weighted finite automaton over \( X \) and \( K \). Then \( \mathcal{A}_N = (A_N, \delta_N, \sigma_N, \tau_N) \) is a crisp-deterministic weighted automaton over \( X \) and \( K \), and it is \( i \)-equivalent to \( \mathcal{A} \), i.e., \( \| \mathcal{A}_N \| = \| \mathcal{A} \| \).

Proof. Let \( u, v \in X^* \) such that \( \sigma_u = \sigma_v \). Then for every \( x \in X \) we have that \( \sigma_{ux} = \sigma_u \cdot \delta_x = \sigma_v \cdot \delta_x = \sigma_{vx} \), and hence, \( \delta_x \) is a well-defined mapping. It is clear that \( \tau_N \) is also a well-defined mapping. Therefore, we have that \( \mathcal{A}_N = (A_N, \delta_N, \sigma_N, \tau_N) \) is a cdwa. According to (10), (13) and (2),

\[
\left(\| \mathcal{A}_N \|, u\right) = \tau_N(\delta_N(\sigma_u, u)) = \tau_N(\sigma_u) = \sigma_u \cdot \tau = \left(\| \mathcal{A} \|, u\right),
\]

for every \( u \in X^* \), and we have proved that \( \mathcal{A}_N \) is \( i \)-equivalent to \( \mathcal{A} \). \( \square \)

Let us note that the cdwa \( \mathcal{A}_N \) constructed above is not necessarily finite. The finiteness of \( \mathcal{A}_N \) depends on the properties of the underlying strong bimonoid \( K \), but not only on that. For determining some necessary and sufficient conditions under which \( \mathcal{A}_N \) is finite we need a notion of local finiteness of a strong bimonoid.

For every subset \( A \) of a strong bimonoid \( K \), the \emph{weak closure} of \( A \), denoted by \( \cl(A) \), is the smallest subset \( B \subseteq K \) such that \( A \subseteq B \) and \( b + b' \in B \) and \( b \cdot a \in B \) for all \( a \in A \) and \( b, b' \in B \). We say that \( K \) is \emph{weakly locally finite} if \( \cl(A) \) is finite for every finite subset \( A \subseteq K \) (cf. [11]).

Namely, the cdwa \( \mathcal{A}_N \) is finite for every wfa \( \mathcal{A} \) over a strong bimonoid \( K \) if \( K \) is weakly locally finite. Obviously, for every strong bimonoid \( K \), if \( \sigma(a) = 0 \), for each \( a \in A \), then \( |A_N| = 1 \). In the sequel we will determine some necessary and sufficient conditions under which the cdwa \( \mathcal{A}_N \) is finite.

By the proof of the previous proposition we obtain the following:

Proposition 6.2. Let \( \mathcal{A} = (A, \delta, \sigma, \tau) \) be a weighted finite automaton over \( X \) and \( K \), and let \( N_\sigma \) be a relation on the free monoid \( X^* \) defined by

\[
(u, v) \in N_\sigma \Leftrightarrow \sigma_u = \sigma_v
\]

for all \( u, v \in X^* \). Then \( N_\sigma \) is a right congruence on \( X^* \).

The right congruence \( N_\sigma \) defined above will be called the \emph{Nerode right congruence} on \( X^* \) determined by \( \mathcal{A} \) and \( \sigma \). or simply, the \emph{Nerode right congruence} of \( \mathcal{A} \). It is clear that there is a bijective correspondence between the set \( A_N \) and the factor set \( X^*/N_\sigma \), and we can regard \( \mathcal{A}_N \) as the right congruence automaton associated with \( N_\sigma \). For that reason we will call \( \mathcal{A}_N \) the \emph{Nerode automaton} of \( \mathcal{A} \) (see [18, 19]).

Let \( \mathcal{A} = (A, \delta, \sigma, \tau) \) be a wfa over \( X \) and \( K \). For any state \( a \in A \) we define an initial weight vector \( \sigma^a : A \rightarrow K \) by \( \sigma^a(a) = 1 \) and \( \sigma^a(b) = 0 \) for every \( b \in A \setminus \{a\} \). The Nerode right congruence determined by the initial variant \( \mathcal{A}^{(a)} = (A, \delta, \sigma^a, \tau) \) of \( \mathcal{A} \) will be denoted by \( N_a \). In this case by Theorem 5.1 we have that \( \sigma^a_u(b) = \delta_u(a, b) \) for all \( u \in X^* \) and \( b \in A \), and \( N_a \) is represented by

\[
(u, v) \in N_a \Leftrightarrow (\forall b \in A) \delta_u(a, b) = \delta_v(a, b),
\]

(15)

for all \( u, v \in X^* \).

For a wfa \( \mathcal{A} = (A, \delta, \sigma, \tau) \) over \( X \) and \( K \) and for every \( a \in A \), let us also define a series \( \| \mathcal{A} \|^{(a)} \) in \( K\langle X^* \rangle \) by

\[
\left(\| \mathcal{A} \|^{(a)} \right), u = \sigma_u(a)
\]

(16)

for every \( u \in X^* \). It is easy to verify that \( \| \mathcal{A} \|^{(a)} = \| \mathcal{A}^{(a)} \|_v \), where \( \mathcal{A}^{(a)} \) is the final variant of \( \mathcal{A} \) given by \( \mathcal{A}^{(a)} = (A, \delta, \sigma, \tau^a) \), and \( \tau^a : A \rightarrow K \) is the final weight vector defined by \( \tau^a(a) = 1 \) and \( \tau^a(b) = 0 \) for every \( b \in A \setminus \{a\} \).

Theorem 6.3. Let \( \mathcal{A} = (A, \delta, \sigma, \tau) \) be a weighted finite automaton over \( X \) and \( K \). Then the following conditions are equivalent:

(i) The Nerode automaton \( \mathcal{A}_N \) is finite.
(ii) The Nerode right congruence \( N_\sigma \) of \( \mathcal{A} \) has finite index.
(iii) The series $\|A\|^{(a,\omega)}$ has a finite image, for every $a \in A$.
(iv) Any series from $\text{Rec}(A, \sigma)$ is cdwfa-recognizable.

Proof. (i) $\Leftrightarrow$ (ii). By the definition of the Nerode right congruence it follows that the rule $\sigma_u \mapsto [u]_{\mathcal{N}_u}$, for every $u \in X^*$, defines a bijective mapping of $\mathcal{N}_u$ onto $X^*/\mathcal{N}_u$, what means that (i) $\Leftrightarrow$ (ii) holds.

(i) $\Rightarrow$ (iv). By Proposition 6.1 we obtain that every final variant $(A, \delta, \sigma, \tau')$ of $\mathcal{A}$ is $i$-equivalent to the final variant $(A_{N_i}, \delta_{N_i}, \sigma_i, \tau_{N_i})$ of $\mathcal{A}_{N_i}$, where the vector $\tau_{N_i} : A_N \rightarrow K$ is defined by $\tau_{N_i}(\sigma_u) = \sigma_u \cdot \tau'$, for each $u \in X^*$. Hence, any series from $\text{Rec}(\mathcal{A}, \sigma)$ is the behavior of some final variant of $\mathcal{A}_{N_i}$, so (i) $\Rightarrow$ (iv) is true.

(iv) $\Rightarrow$ (iii). For every $a \in A$, the series $\|A\|^{(a,\omega)}$ is the $i$-behavior of the final variant $\mathcal{A}^{(a,\omega)}$ of $\mathcal{A}$, and hence, $\|A\|^{(a,\omega)}$ belongs to $\text{Rec}(\mathcal{A}, \sigma)$. This means that $\|A\|^{(a,\omega)}$ is cdwfa-recognizable and therefore has a finite image.

(iii) $\Rightarrow$ (i). For each $u \in X^*$ and $a \in A$ we have $\sigma_u(a) = (\|A\|^{(a,\omega)}, u)$. Hence, the number of different functions $\sigma_u$ ($u \in X^*$) is bounded by $\prod_{a \in A} |\text{Im}(\|A\|^{(a,\omega)})|$ which is finite by assumption. □

Algorithm 6.4 (Construction of the Nerode automaton). The input of this algorithm is a weighted finite automaton $\mathcal{A} = (A, \delta, \sigma, \tau)$ over $X$ and $K$, and the output is the Nerode automaton $\mathcal{A}_N = (A_{N_i}, \delta_{N_i}, \sigma_i, \tau_{N_i})$ which is $i$-equivalent to $\mathcal{A}$.

The procedure is to construct the transition tree of $\mathcal{A}_N$ directly from $\mathcal{A}$. It is constructed inductively in the following way:

(A1) The root of the tree is $\sigma$, and we put $T_0 = \{\sigma\}$.

(A2) After the $i$-th step let a tree $T_i$ have been constructed, and vertices in $T_i$ have been labelled either ‘closed’ or ‘non-closed’. The meaning of these two terms will be made clear in the sequel.

(A3) In the next step we construct a tree $T_{i+1}$ by enriching $T_i$ as follows: for each non-closed leaf $\sigma_a$ occurring in $T_i$, where $u \in X^*$, and each $x \in X$ we add a vertex $\sigma_{ax}$ and an edge from $\sigma_a$ to $\sigma_{ax}$ labelled by $x$. If, in addition, $\sigma_{ax}$ is a state that has already been constructed, then we mark $\sigma_{ax}$ as closed. The procedure terminates when all leaves are marked closed.

(A4) Simultaneously, for each non-closed leaf $\sigma_u$ occurring in $T_i$, where $u \in X^*$, we compute the value $\tau_{N}(\sigma_u)$ using the formula (13).

(A5) When the transition tree of $\mathcal{A}_N$ is constructed, we erase all closure marks and glue leaves to interior vertices with the same label. The diagram that results is the transition graph of $\mathcal{A}_N$.

If the strong bimonoid $K$ is weakly locally finite, then the algorithm terminates in a finite number of steps, for any wfa over $K$, and the result is a cdwfa.

On the other hand, if $K$ is not weakly locally finite, then the algorithm terminates in a finite number of steps under conditions determined by Theorem 6.3.

In view of this, the question arises whether these conditions are decidable. If $K$ is a field (with computable operations) it is decidable whether a wfa over $X$ and $K$ has a finite image, cf. [1, Section VI.2]. This is also true if $K$ is the tropical or the arctic semiring [33]. It would be interesting to obtain such decidability (or undecidability) results for strong bimonoids which are not semirings.

Theorem 6.5. For every weighted finite automaton $\mathcal{A} = (A, \delta, \sigma, \tau)$ over $X$ and $K$, the Nerode automaton $\mathcal{A}_N$ is isomorphic to a final variant of the accessible part of the direct product of derivative automata of series $\|A\|^{(a,\omega)}$, where $a \in A$.

In addition, $\mathcal{A}_N$ is isomorphic to a final variant of a subdirect product of derivative automata of series $\|A\|^{(a,\omega)}$, where $a \in A$.

Proof. Let $A = \{a_1, \ldots, a_n\}$, for some $n \in \mathbb{N}$.

For any $i \in \{1, \ldots, n\}$ let $\psi_i = \|A\|^{(a,\omega)}$, and let $\mathcal{A}_i = (A_i, \delta_i, a_i^0, \tau_i)$ be the derivative automaton of the series $\psi_i$, that is, $A_i = \{u^{-1} \psi_i | u \in X^*\}$, $a_i^0 = \psi_i$, and $\delta_i : A_i \times X \rightarrow A_i$, and $\tau_i : A_i \rightarrow K$ are defined by $\delta_i(\psi, x) = x^{-1} \psi$ and $\tau_i(\psi) = (\psi, \varepsilon)$, for all $\psi \in A_i$ and $x \in X$. 11
Moreover, let $\mathcal{A}_{11} = (A, \delta_{11}, \alpha_0, \tau_{11})$ be the direct product of derivative automata $\mathcal{A}_i$, $i \in \{1, \ldots, n\}$, i.e., $\mathcal{A}_{11} = A_1 \times \cdots \times A_n$, $\alpha_0 = (\varphi_1, \ldots, \varphi_n)$, and $\delta_{11} : A_{11} \times X \to A_{11}$ and $\tau_{11} : A_{11} \to K$ are defined by

$$\delta_{11}(\psi_1, \ldots, \psi_n, x) = (\delta_1(\psi_1, x), \ldots, \delta_n(\psi_n, x)) = (x^{-1}\psi_1, \ldots, x^{-1}\psi_n),$$

$$\tau_{11}(\psi_1, \ldots, \psi_n) = \tau_1(\psi_1) \cdots \tau_n(\psi_n) = \psi_1(\varepsilon) \cdots \psi_n(\varepsilon),$$

for all $(\psi_1, \ldots, \psi_n) \in A_{11}$ and $x \in X$. In fact, $\tau_{11}$ is unessential, because here we will not work with $\mathcal{A}_{11}$, but with a final variant of the accessible part of $\mathcal{A}_{11}$.

Define a mapping $\phi : A_N \to A_{11}$ by $\phi(\sigma_u) = (u^{-1}\varphi_1, \ldots, u^{-1}\varphi_n)$, for every $u \in X^\ast$. For any $u, v \in X^\ast$ we have that

$$\sigma_u = \sigma_v \iff (\forall w \in X^\ast)(\sigma_{uw} = \sigma_{vw})$$

$$= (\forall w \in X^\ast)(\forall i \in \{1, 2, \ldots, n\})(\sigma_{uw}(a_i) = \sigma_{vw}(a_i))$$

$$= (\forall w \in X^\ast)(\forall i \in \{1, 2, \ldots, n\})(\varphi_i(\xi, uw) = \varphi_i(\xi, vw))$$

$$= (\forall i \in \{1, 2, \ldots, n\})(\forall w \in X^\ast)(u^{-1}\varphi_i = v^{-1}\varphi_i)$$

$$= (\forall w \in X^\ast)\phi(\sigma_u = \phi(\sigma_v),$$

and hence, $\phi$ is single-valued and injective. Evidently, $\phi(\sigma) = \phi(\sigma_\varepsilon) = (\varphi_1, \ldots, \varphi_n) = \alpha_0$. Also, for arbitrary $u \in X^\ast$ and $x \in X$ we obtain

$$\phi(\delta_N(\alpha_u, x)) = \phi(\sigma_{ux}) = (u^{-1}\varphi_1, \ldots, u^{-1}\varphi_n)$$

$$= (x^{-1}\varphi_1, \ldots, x^{-1}\varphi_n)$$

$$= \delta_{11}(u^{-1}\varphi_1, \ldots, u^{-1}\varphi_n, x) = \delta_{11}(\phi(\sigma_u), x),$$

whence $\phi(\delta_N(\alpha_u, v)) = \delta_{11}(\phi(\sigma_u), v)$. Finally, for any $u \in X^\ast$ we have

$$\phi(\sigma_u) = (u^{-1}\varphi_1, \ldots, u^{-1}\varphi_n) = \delta_{11}((\varphi_1, \ldots, \varphi_n), u) = \delta_{11}(\alpha_0, u).$$

Now, if $A_{11}' = \text{Im } \phi, \alpha_0' = \alpha_0, \delta_{11}' : A_{11}' \times X \to A_{11}'$ is the restriction of $\delta_{11}$ to $A_{11}'$, and $\tau_{11}' : A_{11} \to K$ is defined by $\tau_{11}'(a) = \tau_N(\phi^{-1}(a))$, for every $a \in A_{11}'$, then $\mathcal{A}_{11}' = (A_{11}', \delta_{11}', \alpha_0', \tau_{11}')$ is a cdwa which is a final variant of the accessible part of $\mathcal{A}_{11}$, and $\phi$ is an isomorphism of $\mathcal{A}_N$ onto $\mathcal{A}_{11}'$.

The final statement is immediate by the first statement and previous remarks, since the derivative automata $\mathcal{A}_i$ are accessible.

\begin{theorem}
Let $\mathcal{A} = (A, \delta, \sigma, \tau)$ be a weighted finite automaton over $X$ and $K$. Then the following holds:

(a) The Nerode automaton $\mathcal{A}_N$ is minimal among all crisp-deterministic weighted automata whose final variants recognize all series $\|A\|^{(\omega)}$, where $a \in A$;

(b) The Nerode automaton $\mathcal{A}_N$ is minimal among all crisp-deterministic weighted automata whose final variants recognize all series from $\text{Rec}(\mathcal{A}, \sigma)$.

\end{theorem}

\begin{proof}
(a) By the proof of Theorem 6.3, any series from $\text{Rec}(\mathcal{A}, \sigma)$ is the behavior of some final variant of $\mathcal{A}_N$, and since $\|A\|^{(\omega)}$ belongs to $\text{Rec}(\mathcal{A}, \sigma)$, for every $a \in A$, we have that it is recognized by some final variant of $\mathcal{A}_N$. Let $\mathcal{A}' = (A', \delta', \sigma', \tau')$ be an arbitrary cdwa such that for every $a \in A$ the series $\|A\|^{(\omega)}$ is recognized by some final variant of $\mathcal{A}'$, i.e., there exists $\tau'_u : A' \to K$ such that $\tau'_u(\delta''(a'_0, u)) = (\|A\|^{(\omega)}, u) = \sigma_u(a)$, for every $u \in X^\ast$. We may assume that $\mathcal{A}'$ is accessible.

Define a mapping $\phi : A' \to A_N$ by setting $\phi(\sigma') = \sigma_u$, where $\sigma' \in A'$ and $u \in X^\ast$ such that $a' = \delta''(a'_0, u)$. To show that $\phi$ is well-defined, consider $a' \in A'$ and $u, v \in X^\ast$ such that $a' = \delta''(a'_0, u) = \delta''(a'_0, v)$. Then for every $a \in A$ we have that

$$\sigma_u(a) = \tau'_u(\delta''(a'_0, u)) = \tau'_u(\delta''(a'_0, v)) = \sigma_v(a),$$

\end{proof}
and hence, \( \sigma_u = \sigma_v \). It is clear that \( \phi \) is a surjective mapping, so we conclude that \(|A_N| \leq |A'|\). Therefore, we have proved that \( A_N \) is a minimal cdwa whose final variants recognize all series \( \|A\|_{(a,b)}, a \in A \).

(b) This follows by (a) and the fact that all series \( \|A\|_{(a,b)}, a \in A \), belong to \( \text{Rec}(A, \sigma) \). □

7. Determinization related to the transition semantics

Let \( \mathcal{A} = (A, \delta, \sigma, \tau) \) be a wfa over \( X \) and \( K \). Let us set \( A_M = \{ \delta_u \mid u \in X^* \} \), and let us define \( \delta_M : A_M \times X \to A_M \) and \( \tau_M : A_M \to K \) by

\[
\delta_M(\delta_u, x) = \delta_{ux} \quad \text{and} \quad \tau_M(\delta_u) = (\sigma \cdot \delta_u) \cdot \tau,
\]

for every \( u \in X^* \).

**Proposition 7.1.** Let \( \mathcal{A} = (A, \delta, \sigma, \tau) \) be a weighted finite automaton over \( X \) and \( K \). Then \( \mathcal{A}_M = (A_M, \delta_M, \delta, \tau_M) \) is a crisp-deterministic weighted automaton over \( X \) and \( K \), and it is \( t \)-equivalent to \( \mathcal{A} \), i.e., \( \|A_M\| = \|A\| \).

**Proof.** Let \( u, v \in X^* \) such that \( \delta_u = \delta_v \). Then for every \( x \in X \) we have that \( \delta_{ux} = \delta_u \cdot \delta_x = \delta_v \cdot \delta_x = \delta_{vx} \), and hence, \( \delta_M \) is a well-defined mapping. It is clear that \( \tau_M \) is also a well-defined mapping. Therefore, we have that \( \mathcal{A}_M = (A_M, \delta_M, \delta, \tau_M) \) is a cdwa. According to (10), (17) and (4),

\[
(\|A_M\|, u) = \tau_M(\delta_M(\delta, u)) = \tau_M(\delta_u) = (\sigma \cdot \delta_u) \cdot \tau = (\|A\|, u),
\]

for every \( u \in X^* \), and we have proved that \( \mathcal{A}_M \) is \( t \)-equivalent to \( \mathcal{A} \). □

**Proposition 7.2.** Let \( \mathcal{A} = (A, \delta, \sigma, \tau) \) be a weighted finite automaton over \( X \) and \( K \), and let \( M_{\mathcal{A}} \) be the relation on the free monoid \( X^* \) defined by

\[
(u, v) \in M_{\mathcal{A}} \iff \delta_u = \delta_v,
\]

for all \( u, v \in X^* \). Then \( M_{\mathcal{A}} \) is a right congruence on \( X^* \).

**Proof.** Clear, since \( \delta_{ux} = \delta_u \cdot \delta_x \) for all \( u \in X^* \) and \( x \in X \). □

The congruence \( M_{\mathcal{A}} \) defined above will be called the Myhill right congruence on \( X^* \) determined by \( \mathcal{A} \). Let \( M = M_{\mathcal{A}} \) and let \([u]_M\) denote the equivalence class of \( u \in X^* \) with respect to \( M \). Since the rule \( \delta_u \mapsto [u]_M \), for each \( u \in X^* \), defines a bijective mapping of \( A_M \) onto \( X^* / M_{\mathcal{A}} \), we can regard the cdwa \( \mathcal{A}_M \) as the right congruence automaton associated with \( M_{\mathcal{A}} \), and the cdwa \( \mathcal{A}_M \) is called the Myhill automaton of \( \mathcal{A} \) (see [19]).

**Example 7.3.** Let \( K = ([0,1], \oplus, \cdot, 0, 1) \) be the strong bimonoid given in Example 2.1 (3), where \( \oplus \) is the bounded sum, and let \( \mathcal{A} \) be any wfa over an alphabet \( X = \{x, y, z\} \) and \( K \) with weighted transition matrices

\[
\delta_x = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \quad \delta_y = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \quad \delta_z = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 0 \end{bmatrix}.
\]

Then \( \delta_{xy} = \delta_x \cdot \delta_y = \delta_x \cdot \delta_y = \delta_{xy} \), so \( (zy, xy) \in M_{\mathcal{A}} \), but \( \delta_{zyy} = (\delta_z \cdot \delta_y) \cdot \delta_y \neq (\delta_z \cdot \delta_z) \cdot \delta_y = \delta_{zyy} \), whence \( (zy, zxy) \notin M_{\mathcal{A}} \). Thus, \( M_{\mathcal{A}} \) is not a left congruence. We also have that \( \delta_z \cdot (\delta_z \cdot \delta_y) \neq (\delta_z \cdot \delta_z) \cdot \delta_y \), and hence, weighted transition matrices do not form a semigroup.

For a wfa \( \mathcal{A} = (A, \delta, \sigma, \tau) \) over \( X \) and \( K \), and for every \( (a, b) \in A \times A \), we define a series \( \|A\|_{(a,b)} \) in \( K\langle X^* \rangle \) by

\[
(\|A\|_{(a,b)}, u) = \delta_{ub}(a, b)
\]

for every \( u \in X^* \). It is easy to verify that \( \|A\|_{(a,b)} = \|A\|_{(a,b)} \), where the wfa \( A_{(a,b)} \) is the variant of \( A \) given by \( A_{(a,b)} = (A, \delta, \sigma^a, \tau^b) \), and the vectors \( \sigma^a, \tau^b : A \to K \) are defined by \( \sigma^a(a) = \tau^b(b) = 1 \) and \( \sigma^a(c) = 0 \) and \( \tau^b(d) = 0 \) for all \( c \in A \setminus \{a\} \) and \( d \in A \setminus \{b\} \).
Theorem 7.4. Let $A = (A, \delta, \sigma, \tau)$ be a weighted finite automaton over $X$ and $K$. Then the following conditions are equivalent:

(i) The Myhill automaton $A_M$ is finite.
(ii) The Myhill right congruence $M_A$ of $A$ has finite index.
(iii) The Nerode right congruence $N_a$ has finite index, for every $a \in A$.
(iv) The series $\|A\|^{a,b}$ has a finite image, for every $(a, b) \in A \times A$.
(v) Any series from $\text{Rec}(A)$ is cdwfa-recognizable.

Proof. (i)$\Rightarrow$(ii). As we have noted above, there is a bijective mapping of $A_M$ onto $X'/M_A$, what means that (i)$\Rightarrow$(ii) holds.

(ii)$\Rightarrow$(iii). According to (15), for every $a \in A$ we have that $M_A \subseteq N_a$, what implies
\[ \text{ind}(N_a) \leq \text{ind}(M_A) < \infty. \]
by the hypothesis.

(iii)$\Rightarrow$(ii). According to (18) and (15) we have that
\[ M_A = \bigcap_{a \in A} N_a. \]
Since $A$ is finite and $N_a$ has finite index for every $a \in A$, it follows that
\[ \text{ind}(M_A) \leq \prod_{a \in A} \text{ind}(N_a) < \infty. \]

(i)$\Rightarrow$(v). This follows immediately by Proposition 7.1.
(v)$\Rightarrow$(iv). Since $\|A\|^{a,b} = \|A(a,b)\|$, and $A(a,b)$ is a variant of $A$, by (v) it follows that the series $\|A\|^{a,b}$ is cdwfa-recognizable and hence has a finite image.

(iv)$\Rightarrow$(iii). Choose $a \in A$. By assumption, the set $\{\delta_u(a, b) \mid u \in X^*, b \in A\}$ is finite. By (15), it follows that $N_a$ has finite index. \hfill \qed

Next, we describe an algorithm for constructing the Myhill automaton. It is very similar to the construction of the Nerode automaton. We just replace the state set $A_N$ by $A_M$ and the root of the tree $\sigma_\varepsilon$ by $\delta_\varepsilon$.

Algorithm 7.5 (Construction of the Myhill automaton). The input of this algorithm is a weighted finite automaton $A = (A, \delta, \sigma, \tau)$ over $X$ and $K$, and the output is the Myhill automaton $A_M = (A_M, \delta_M, \sigma_M, \tau_M)$ which is $t$-equivalent to $A$.

The procedure is to construct the transition tree of $A_M$ directly from $A$. This is done in the same way as in Algorithm 6.4 except that we replace: in (A1): $\sigma_\varepsilon$ by $\delta_\varepsilon$; in (A3) and (A4): $\sigma_u$ and $\sigma_{ux}$ by, resp., $\delta_u$ and $\delta_{ux}$; in (A4): $\tau_N$ by $\tau_M$, and (13) by (17); in (A5): $A_N$ by $A_M$.

If the strong bimonoid $K$ is weakly locally finite, then the algorithm terminates in a finite number of steps, for any wfa over $K$, and the result is a cdwfa.

On the other hand, if $K$ is not weakly locally finite, then the algorithm terminates in a finite number of steps under conditions determined by Theorem 7.4.

Theorem 7.6. For every weighted finite automaton $A = (A, \delta, \sigma, \tau)$ over $X$ and $K$, the Myhill automaton $A_M$ is isomorphic to a final variant of the accessible part of the direct product of derivative automata of series $\|A\|^{a,b}$, where $(a, b) \in A^2$.

In addition, $A_M$ is isomorphic to a final variant of a subdirect product of derivative automata of series $\|A\|^{a,b}$, where $(a, b) \in A^2$.

Proof. The proof of this theorem is similar to the proof of Theorem 6.5 and it will be omitted. \hfill \qed
**Theorem 7.7.** For every weighted finite automaton $\mathcal{A} = (A, \delta, \sigma, \tau)$ over $X$ and $K$, the Myhill automaton $\mathcal{A}_M$ is isomorphic to a final variant of the accessible part of the direct product of Nerode automata $\mathcal{A}_{\mathcal{N}_a}$, where $a \in A$.

In addition, $\mathcal{A}_M$ is isomorphic to a final variant of a subdirect product of Nerode automata $\mathcal{A}_{\mathcal{N}_a}$, where $a \in A$.

**Proof.** Let $A = \{a_1, \ldots, a_n\}$, for some $n \in \mathbb{N}\setminus\{0\}$.

For any $i \in \{1, \ldots, n\}$ set

$$A_i = A_{\mathcal{N}_a} = \{a_i^u \mid u \in X^i\}, \quad \delta_i = \delta_{\mathcal{N}_a}, \quad \sigma_i = \sigma_{\mathcal{N}_a}, \quad \tau_i = \tau_{\mathcal{N}_a}, \quad \mathcal{A}_i = (A_i, \delta_i, \sigma_i, \tau_i).$$

Let $\mathcal{A}_{11} = (A, \delta_{11}, \alpha_0, \tau_{11})$ be the direct product of automata $\mathcal{A}_i$, $i \in \{1, \ldots, n\}$, i.e., $\mathcal{A}_{11} = A_1 \times \cdots \times A_n$, $\alpha_0 = (\sigma_{\alpha_0}^1, \ldots, \sigma_{\alpha_0}^n)$, and $\tau_{11} : \mathcal{A}_{11} \times X \rightarrow \mathcal{A}_{11}$ and $\tau_{11} : \mathcal{A}_{11} \rightarrow K$ are defined by

$$\delta_{11}(\sigma_{\alpha_0}^1, \ldots, \sigma_{\alpha_0}^n, x) = \left(\delta_1(\sigma_{\alpha_0}^1, x), \ldots, \delta_n(\sigma_{\alpha_0}^n, x)\right) = (\sigma_{\alpha_1}^1, \ldots, \sigma_{\alpha_n}^n),$$

$$\tau_{11}(\sigma_{\alpha_0}^1, \ldots, \sigma_{\alpha_0}^n) = \tau_1(\sigma_{\alpha_0}^1) \cdots \tau_n(\sigma_{\alpha_0}^n),$$

for every $(\sigma_{\alpha_0}^1, \ldots, \sigma_{\alpha_0}^n) \in \mathcal{A}_{11}$, where $u_1, \ldots, u_n \in X^i$, and every $x \in X$. Observe that $\sigma_{u_i}^n = \sigma_{\alpha_i}^n \cdot \delta_{\alpha_i}$, by Theorem 5.1.

Define a mapping $\phi : A_M \rightarrow \mathcal{A}_{11}$ by $\phi(\delta_u) = (\sigma_{\alpha_0}^1, \ldots, \sigma_{\alpha_0}^n)$, for every $u \in X^i$. For any $u, v \in X^i$ we have

$$\delta_u = \delta_v \iff (\forall i, j \in \{1, 2, \ldots, n\}) \delta_u(a_i, a_j) = \delta_v(a_i, a_j)$$

$$\iff (\forall i, j \in \{1, 2, \ldots, n\}) (\sigma_{\alpha_0}^i \cdot \delta_u)(a_i) = (\sigma_{\alpha_0}^i \cdot \delta_v)(a_i)$$

$$\iff (\forall i, j \in \{1, 2, \ldots, n\}) \sigma_{\alpha_0}^i(a_i) = \sigma_{\alpha_0}^i(a_i)$$

$$\iff (\forall i \in \{1, 2, \ldots, n\}) \sigma_{u_i}^n = \sigma_{v_i}^n$$

$$\iff \phi(\delta_u) = \phi(\delta_v),$$

and hence, $\phi$ is single-valued and injective. We have that $\phi(\delta_u) = (\sigma_{\alpha_0}^1, \ldots, \sigma_{\alpha_0}^n) = \alpha_0$. Also, for all $u \in X^i$ and $x \in X$ we obtain

$$\phi(\delta_{M\mathcal{N}_a}(\delta_u, x)) = \phi(\delta_{\mathcal{N}_a}) = (\sigma_{\alpha_0}^1, \ldots, \sigma_{\alpha_0}^n) = \delta_{11}(\sigma_{\alpha_0}^1, \ldots, \sigma_{\alpha_0}^n, x) = \delta_{11}(\phi(\delta_u), x).$$

Finally, for any $u \in X^i$ we have

$$\phi(\sigma_u) = (\sigma_{\alpha_0}^1, \ldots, \sigma_{\alpha_0}^n) = \delta_{11}(\sigma_{\alpha_0}^1, \ldots, \sigma_{\alpha_0}^n, u) = \delta_{11}(\alpha_0, u).$$

Now, if $\mathcal{A}_{11}' = \text{Im } \phi$, $\alpha_0' = \alpha_0$, $\delta_{11}' : \mathcal{A}_{11}' \times X \rightarrow \mathcal{A}_{11}'$ is the restriction of $\delta_{11}$ to $\mathcal{A}_{11}'$, and $\tau_{11}' : \mathcal{A}_{11}' \rightarrow K$ is defined by $\tau_{11}'(\alpha) = \tau_M(\phi^{-1}(\alpha))$, for every $\alpha \in \mathcal{A}_{11}'$, then $\mathcal{A}_{11}' = (\mathcal{A}_{11}', \delta_{11}', \alpha_0', \tau_{11}')$ is a cdwa which is a final variant of the accessible part of $\mathcal{A}_{11}$, and $\phi$ is an isomorphism of $\mathcal{A}_M$ onto $\mathcal{A}_{11}'$.

The final statement is immediate by the first statement and previous remarks, since the Nerode automata $\mathcal{A}_{\mathcal{N}_a}$ are accessible. \qed

**Theorem 7.8.** Let $\mathcal{A} = (A, \delta, \sigma, \tau)$ be a weighted finite automaton over $X$ and $K$. Then

(a) The Myhill automaton $\mathcal{A}_M$ is minimal among all crisp-deterministic weighted finite automata whose final variants recognize all series $\llbracket \mathcal{A} \rrbracket_{\alpha(b)}$, where $(a, b) \in A^2$;

(b) The Myhill automaton $\mathcal{A}_M$ is minimal among all crisp-deterministic weighted finite automata whose final variants recognize all series from $\text{Rec}_C(\mathcal{A})$.

**Proof.** The proof of this theorem is similar to the proof of Theorem 6.6 and it will be omitted. \qed
8. Determinization related to the run semantics: The multiset construction

Let $\mathcal{A} = (A, \delta, \sigma, \tau)$ be a wfa over $X$ and $K$. Moreover, let $T$ be the submonoid of $(K, \cdot, 1)$ generated by $\text{Im}(0)$, and let $R = (\text{Im}(a) \cdot T) \setminus \{0\}$. Suppose that $R$ is finite, and that every element $k \in R \cdot \text{Im}(\tau)$ has a finite additive order. So, $A \times R$ is finite. Set $i = \max \{i(k) \mid k \in R \cdot \text{Im}(\tau)\}$ and $p = \text{lcm}[p(k) \mid k \in R \cdot \text{Im}(\tau)]$ (the least common multiple), where $i(k)$ is the index of $k$ and $p(k)$ is the period of $k$ in $(K, +)$. In particular, if $K$ is bi-locally finite, then every wfa over $X$ and $K$ has the above properties.

For every $(a, k) \in A \times K$ and every $u = x_1x_2\cdots x_n \in X^*$, where $x_1, x_2, \ldots, x_n \in X$, $n \in \mathbb{N}$, let us define a set $P_u(a, k)$ by

$$P_u(a, k) = \{(a_0, a_1, \ldots, a_{n-1}, a_n) \in A^{n+1} \mid \sigma(a_0) \cdot \delta_{x_1}(a_0, a_1) \cdot \delta_{x_2}(a_1, a_2) \cdot \delta_{x_3}(a_2, a_3) \cdots \delta_{x_n}(a_{n-1}, a_n) = k, a_n = a\};$$

for $u = \epsilon$ and $n = 0$ this includes the case

$$P_\epsilon(a, k) = \begin{cases} \{a\} & \text{if } k = \sigma(a), \\ \emptyset & \text{if } k \neq \sigma(a). \end{cases}$$

In the graph representation of $\mathcal{A}$, we can regard $P_u(a, k)$ as the set of all paths labeled by $u$, ending in $a$, of weight $k$.

Furthermore, for every $u \in X^*$, let us define $p_u : A \times R \rightarrow \mathbb{N}$ and $\pi_u : A \times R \rightarrow \{0, 1, \ldots, i + p - 1\}$ by $p_u(a, k) = |P_u(a, k)|$ and

$$\pi_u(a, k) = \begin{cases} p_u(a, k) & \text{if } p_u(a, k) < i, \\ i + \left[\left(p_u(a, k) - i\right) \mod p\right] & \text{if } p_u(a, k) \geq i, \end{cases}$$

where $\left(p_u(a, k) - i\right) \mod p$ is the remainder when $p_u(a, k) - i$ is divided by $p$. In the first case $\pi_u(a, k) < i$, and in the second case $i \leq \pi_u(a, k) \leq i + p - 1$, so $\pi_u$ is well-defined. In either case $\pi_u(a, k) \equiv_p p_u(a, k)$, where $\equiv_p$ denotes the congruence modulo $p$. In fact, $p_u(a, k)$ is the number of paths labelled by $u$, ending in $a$, of weight $k$. Let us note that $p_u$ is a multiset, and $\pi_u$ can be considered as some kind of a multiset with a reduced multiplicity function (see [30], as well as other chapters of [6]).

First we prove that the following is true:

Lemma 8.1. Let $\mathcal{A} = (A, \delta, \sigma, \tau)$ be a weighted finite automaton over $X$ and $K$. Then for all $(a, k) \in A \times R$, $u \in X^*$ and $x \in X$ we have that

$$p_{ux}(a, k) = \sum_{(b, l) \in Q_{ux}(a, k)} p_u(b, l),$$

where

$$Q_{ux}(a, k) = \{(b, l) \in A \times R \mid \pi_u(b, l) > 0 \text{ and } l \cdot \delta_x(b, a) = k\}.$$ (24)

Proof. For each $(b, l) \in A \times R$ let us define the set $P_u'(b, l)$ by

$$P_u'(b, l) = \{(b_0, b_1, \ldots, b_{n-1}, b_n, a) \in A^{n+2} \mid (b_0, b_1, \ldots, b_{n-1}, b_n) \in P_u(b, l)\}.$$ (25)

Clearly, $|P_u'(b, l)| = |P_u(b, l)| = p_u(b, l)$. Note that $P_u'(b, l) \cap P_u'(c, j) = \emptyset$, for all $(b, l), (c, j) \in A \times R$ with $(b, l) \neq (c, j)$, and

$$p_{ux}(a, k) = \bigcup_{(b, l) \in Q_{ux}(a, k)} p_u'(b, l).$$ (26)

This implies (23). \qed
We note that the condition $\pi_u(b, l) > 0$ in the definition of the set $Q_{ux}(a, k)$ is used in Algorithm 8.4 given below where it reduces the number of elements $\pi_{ux}(a, k)$ to be computed.

Next, let $A_n = \{\pi_u \mid u \in X^*\}$, let $\delta_n : A_n \times X \rightarrow A_n$ be defined by

$$\delta_n(\pi_u, x) = \pi_{ux},$$

(27)

for every $\pi_u \in A_n$ and $x \in X$, and let $\tau_n : A_n \rightarrow K$ be defined by

$$\tau_n(\pi_u) = \sum_{(a, k) \in A_n \times R} \pi_u(a, k)(k \cdot \tau(a)),$$

(28)

for every $\pi_u \in A_n$, where for every $n \in \mathbb{N}\setminus\{0\}$ and $k \in K$ we define $nk = k + \cdots + k$ ($n$ summands), and $0k = 0$. Clearly, $\tau_n$ is well-defined, and we show that $\delta_n$ is well-defined as part of the following theorem:

**Theorem 8.2.** Let $A = (A, \delta, \sigma, \tau)$ be a weighted finite automaton over $X$ and $K$, having the property that the set $R = (\text{Im}(\sigma) \cdot T) \setminus \{0\}$ is finite and every element of $R \cdot \text{Im}(\tau)$ has a finite additive order, where $T$ denotes the submonoid of $(K, \cdot, 1)$ generated by the set $\text{Im}(\delta)$.

Then $A_r = (A_n, \delta_n, \pi_{ux}, \tau_n)$ is an accessible crisp-deterministic weighted finite automaton over $X$ and $K$ which is $r$-equivalent to $A$.

**Proof.** First we show that $\delta_n$ is well-defined. Let $u, v \in X^*$ such that $\pi_u = \pi_v$ and let $(a, k) \in A \times R$ and $x \in X$. By (24) we obtain that $Q_{ux}(a, k) = Q_{vx}(a, k)$. For sake of simplicity, set $Q_{ux}(a, k) = Q_{vx}(a, k) = Q$. Then by Lemma 8.1 we obtain that

$$p_{ux}(a, k) = \sum_{(b, l) \in Q} p_u(b, l) \text{ and } p_{vx}(a, k) = \sum_{(b, l) \in Q} p_v(b, l).$$

(29)

Next, suppose that $\pi_{ux}(a, k) < i$. Then for every $(b, l) \in Q$ we have that

$$p_u(b, l) \leq \sum_{(c, j) \in Q} p_u(c, j) = p_{ux}(a, k) = \pi_{ux}(a, k) < i,$$

so $p_u(b, l) = \pi_u(b, l) = p_v(b, l)$, and hence, $p_u(b, l) = p_v(b, l) = p_u(b, l)$. Therefore,

$$p_{vx}(a, k) = \sum_{(b, l) \in Q} p_v(b, l) = \sum_{(b, l) \in Q} p_u(b, l) = p_{ux}(a, k),$$

which yields

$$\pi_{ux}(a, k) = p_{ux}(a, k) = p_{vx}(a, k) = \pi_{vx}(a, k).$$

(30)

Similarly we prove that $\pi_{vx}(a, k) < i$ implies (30).

Assume now that $i \leq \pi_{ux}(a, k) \leq i + p - 1$ and $i \leq \pi_{vx}(a, k) \leq i + p - 1$. For any $(b, l) \in Q$ we have $p_u(b, l) \equiv_p \pi_u(b, l) = \pi_v(b, l) \equiv_p p_v(b, l)$. Therefore,

$$\pi_{ux}(a, k) \equiv_p p_{ux}(a, k) = \sum_{(b, l) \in Q} p_u(b, l) \equiv_p \sum_{(b, l) \in Q} p_v(b, l) = p_{vx}(a, k) \equiv_p \pi_{vx}(a, k),$$

and since $i \leq \pi_{ux}(a, k), \pi_{vx}(a, k) \leq i + p - 1$, we conclude that $\pi_{ux}(a, k) = \pi_{vx}(a, k)$. Thus $\delta_n$ is well-defined. Consequently, $A_r = (A_n, \delta_n, \pi_{ux}, \tau_n)$ is an accessible crisp-deterministic weighted automaton. Clearly, $A_n$ is finite and $|A_n| \leq (i + p)^{|A| |R|}$

To prove that $A_r$ is $r$-equivalent to $A$, consider any $u = x_1x_2 \cdots x_n \in X^*$, where $x_1, x_2, \ldots, x_n \in X$, $a \in A$ and $k \in R$. As we have already noticed, $\pi_u(a, k) = p_u(a, k)$ or $\pi_u(a, k) \equiv_p p_u(a, k)$, and in the second case $p(k \cdot \tau(a))$ divides $p_u(a, k) - \pi_u(a, k)$. Hence
Therefore, 

\[
\pi_u(a, k)(k \cdot \tau(a)) = p_u(a, k)(k \cdot \tau(a)) = \sum_{(a_0, \ldots, a_n, a) \in P_u(a, k)} \sigma(a_0) \cdot \delta_{x_1}(a_0, a_1) \cdots \delta_{x_n}(a_{n-1}, a_n) \cdot \tau(a).
\]

Now we have that

\[
([A_n], u) = \tau_n(\delta_n^*(\pi_c, u)) = \tau_n(\pi_u) = \sum_{(a, k) \in A \times R} \pi_u(a, k)(k \cdot \tau(a))
\]

\[
= \sum_{(a, k) \in A \times R} \left( \sum_{(a_0, \ldots, a_n, a) \in P_u(a, k)} \sigma(a_0) \cdot \delta_{x_1}(a_0, a_1) \cdots \delta_{x_n}(a_{n-1}, a_n) \cdot \tau(a) \right)
\]

\[
= \sum_{(a_0, \ldots, a_n, a) \in A^{n+1}} \sigma(a_0) \cdot \delta_{x_1}(a_0, a_1) \cdots \delta_{x_n}(a_{n-1}, a) \cdot \tau(a)
\]

\[
= ([A], u).
\]

Therefore, \( A_n \) is \( r \)-equivalent to \( A \). 

Let \( A = (A, \delta, \sigma, \tau) \) be a wfa over \( X \) and \( K \) such that \( \delta \) is crisp. In the notation introduced at the beginning of this section we have that \( T = \text{Im}(\delta) = [0, 1] \) and \( R = \text{Im}(\sigma) \setminus [0] \). This means that the set \( A \times R \) is finite, and if every element of \( \text{Im}(\sigma) \cdot \text{Im}(\tau) \) has a finite additive order, then \( A_n \) is finite.

We note that given the situation of Theorem 8.2, by the proofs of [11] we also obtain a construction of a cdwfa which is \( r \)-equivalent to \( A \). However, the present construction is direct and therefore yields a wfa with a much smaller state set. For instance, assume that \( K \) is a bounded lattice. Then \( i = p = 1 \), so \( |A_n| \leq 2^{|A|R} \). Since \( A_n \) is crisp-deterministic but \( A \) was arbitrary, the exponential blow-up of \( |A| \) to \( 2^{|A|} \) cannot be avoided. Also, again since \( A_n \) is crisp-deterministic, \( A_n \) can realize the values of \([A]\), only by its final weights, therefore \( |A_n| \) has to be exponential in \( |R| \). In view of this, the construction of the state space of \( A_n \) seems to be optimal.

Algorithm 8.3 (Construction of the cdwfa \( A_n \)). The input of this algorithm is a weighted finite automaton \( A = (A, \delta, \sigma, \tau) \) over \( X \) and \( K \), satisfying the properties of Theorem 8.2. The output is the accessible crisp-deterministic weighted finite automaton \( A_n = (A_n, \delta_n, \tau_n, \tau_n) \) which is \( r \)-equivalent to \( A \).

The procedure is to construct the transition tree of \( A_n \) directly from \( A \). This is done in the same way as in Algorithm 6.4 except that we replace: in (A1): \( \sigma_i \) by \( \pi_i \); in (A3) and (A4): \( \sigma_u \) and \( \sigma_{ux} \) by, resp., \( \pi_u \) and \( \pi_{ux} \); in (A4): \( \tau_N \) by \( \tau_u \), and (13) by (28); in (A5): \( \mathcal{A}_y \) by \( \mathcal{A}_n \).

Since \( A \) has the mentioned property that the set \( R = (\text{Im}(\sigma) \cdot T) \setminus [0] \) is finite and every element of \( R \cdot \text{Im}(\tau) \) has a finite additive order, where \( T \) denotes the submonoid of \( (K, \cdot, 1) \) generated by \( \text{Im}(\delta) \), the algorithm terminates in a finite number of steps. In particular, if \( K \) is bi-locally finite, then every wfa over \( X \) and \( K \) has this property. On the other hand, if \( K \) is not bi-locally finite, then the algorithm does not necessarily terminate in a finite number of steps, but it can happen in some cases.

Algorithm 8.4 (Computing \( \pi_{ux} \) from \( \pi_u \)). This procedure computes \( \pi_{ux} \) directly from \( \pi_u \) for every \( u \in X^* \) and \( x \in X \):

1. Initialization: For each \((a, k) \in A \times R \) we set \( \pi_{ux}(a, k) = 0 \).
2. For all \((a, k) \in A \times R \) and \( b \in A \) such that \( \pi_u(a, k) > 0 \) and \( k \cdot \delta_x(a, b) \neq 0 \) we do the following:
   - If \( \pi_{ux}(b, k \cdot \delta_x(a, b)) + \pi_u(a, k) \leq i + p - 1 \), then we change the value \( \pi_{ux}(b, k \cdot \delta_x(a, b)) \) by
     \[
     \pi_{ux}(b, k \cdot \delta_x(a, b)) = \pi_{ux}(b, k \cdot \delta_x(a, b)) + \pi_u(a, k),
     \]
   - otherwise we change the value \( \pi_{ux}(b, k \cdot \delta_x(a, b)) \) by
     \[
     \pi_{ux}(b, k \cdot \delta_x(a, b)) = i + \left[ \pi_{ux}(b, k \cdot \delta_x(a, b)) + \pi_u(a, k) - i \right] \mod p
     \]

By Lemma 8.1, at the end of this procedure we obtain correct values \( \pi_{ux}(a, k) \) for all \((a, k) \in A \times R \).
9. Determinization related to the run semantics: The set construction

In this section, we present a simplification of Theorem 8.2 and Algorithms 8.3 and 8.4 under the assumption that certain elements arising from the weights of the wfa under consideration are additively idempotent. This assumption is automatically satisfied if $K$ is additively idempotent; the ubiquity of this important case is the reason for the development of these simplified algorithms.

Let $\mathcal{A} = (A, \delta, \sigma, \tau)$ be a wfa over $X$ and $K$, let $T$ be the submonoid of $(K, +, 1)$ generated by the set $\text{Im}(\delta)$, and let $R = (\text{Im}(\delta) \cdot T) \setminus \{0\}$. Suppose that $R$ is finite, and that every element $k \in R \cdot \text{Im}(\tau)$ is additively idempotent. Then $i(k) = p(k) = 1$, for every $k \in R \cdot \text{Im}(\tau)$, and in the notation from Section 8, we have that $i = p = 1$. In particular, if $K$ is additively idempotent and multiplicatively locally finite, then every wfa over $X$ and $K$ has the above properties.

Therefore, for every $u = x_1x_2 \cdots x_n \in X^+$, where $x_1, x_2, \ldots, x_n \in X$, $n \in \mathbb{N}\setminus\{0\}$, and every $(a, k) \in A \times R$ we have that

$$
\pi_u(a, k) = \begin{cases} 
0 & \text{if } p_u(a, k) = 0, \\
1 & \text{if } p_u(a, k) \geq 1,
\end{cases}
$$

(31)

and also, for every $(a, k) \in A \times R$ we have that

$$
\pi_\epsilon(a, k) = \begin{cases} 
0 & \text{if } \sigma(a) \neq k, \\
1 & \text{if } \sigma(a) = k.
\end{cases}
$$

(32)

This means that for every $u = x_1x_2 \cdots x_n \in X^+$, where $x_1, x_2, \ldots, x_n \in X$, $n \in \mathbb{N}\setminus\{0\}$, we can regard $\pi_u$ as identical with a set $\varrho_u \subseteq A \times R$ given by

$$
\varrho_u = \{(a, k) \in A \times R \mid \pi_u(a, k) = 1\} = \{(a, k) \in A \times R \mid p_u(a, k) \geq 1\}
$$

$$
= \{(a, \sigma(a) \cdot \delta_{x_1}(a_0, a_1) \cdots \delta_{x_n}(a_{n-1}, a)) \mid (a_0, a_1, \ldots, a_{n-1}, a) \in A^{n+1}, \sigma(a) \cdot \delta_{x_1}(a_0, a_1) \cdots \delta_{x_n}(a_{n-1}, a) \neq 0\},
$$

(33)

and $\pi_\epsilon$ can be regarded as identical with a set $\varrho_\epsilon \subseteq A \times R$ given by

$$
\varrho_\epsilon = \{(a, k) \in A \times R \mid \pi_\epsilon(a, k) = 1\} = \{(a, k) \in A \times R \mid p_\epsilon(a, k) \geq 1\}
$$

$$
= \{(a, \sigma(a)) \mid a \in A, \sigma(a) \neq 0\}.
$$

(34)

According to (33), for every $x \in X$ and $u = x_1x_2 \cdots x_n \in X^+$, where $x_1, x_2, \ldots, x_n \in X$, $n \in \mathbb{N}\setminus\{0\}$, we have that

$$
\varrho_{ux} = \{(b, \sigma(a) \cdot \delta_{x_1}(a_0, a_1) \cdots \delta_{x_n}(a_{n-1}, a) \cdot \delta_x(a, b)) \mid (a_0, a_1, \ldots, a_{n-1}, a, b) \in A^{n+2}, \sigma(a) \cdot \delta_{x_1}(a_0, a_1) \cdots \delta_{x_n}(a_{n-1}, a) \cdot \delta_x(a, b) \neq 0\}
$$

$$
= \{(b, k \cdot \delta_x(a, b)) \mid (a, k) \in \varrho_u, \ b \in A, \ k \cdot \delta_x(a, b) \neq 0\}.
$$

(35)

Clearly, (35) is also true for $u = \epsilon$.

Now, let $A_\varrho = \{\varrho_u \mid u \in X^+\}$, let $\delta_\varrho : A_\varrho \times X \to A_\varrho$ be defined by

$$
\delta_\varrho(\varrho_u, x) = \varrho_{ux},
$$

(36)

for every $\varrho_u \in A_\varrho$ and $x \in X$, and let $\tau_\varrho : A_\varrho \to K$ be defined by

$$
\tau_\varrho(\varrho_u) = \sum_{(a, k) \in \varrho_u} k \cdot \tau(a),
$$

(37)

for every $\varrho_u \in A_\varrho$. Then the following is true:
Theorem 9.1. Let $\mathcal{A} = (A, \delta, \sigma, \tau)$ be a weighted finite automaton over $X$ and $K$, having the property that the set $R = (\text{Im}(\sigma) \cdot T) \setminus \{0\}$ is finite and every element of $R \cdot \text{Im}(\tau)$ is additively idempotent, where $T$ denotes the submonoid of $(K, \cdot, 1)$ generated by the set $\text{Im}(\delta)$.

Then $\mathcal{A}_\phi = (A_\delta, \delta_\phi, \phi_\tau, \tau_\phi)$ is an accessible crisp-deterministic weighted finite automaton over $X$ and $K$, and it is $r$-equivalent to $\mathcal{A}$.

Proof. It is evident that $\tau_\phi$ is well-defined. According to (35), if $\phi_u = \phi_v$, for some $u, v \in X'$, then $\phi_u = \phi_v$, for every $x \in X$, and we have that $\delta_\phi$ is also well-defined. Therefore, $\mathcal{A}_\phi = (A_\delta, \delta_\phi, \phi_\tau, \tau_\phi)$ is an accessible crisp-deterministic weighted automaton. By the hypothesis, $T$ is finite, so $A \times T$ and $A_\delta \subseteq 2^{A \times R}$ are also finite. Consequently, $\mathcal{A}_\phi$ is finite and $|\mathcal{A}_\phi| \leq 2^{|A| \cdot |R|}$.

Next, according to (31), (32), (33), (34) and Theorem 8.2, for every $u \in X'$ we have that

$$\left(\left[\mathcal{A}_\phi\right], u\right) = \tau_\phi(\delta_\phi^*(u, u)) = \tau_\phi(\delta_u) = \sum_{(a,k)\in \delta_u} k \cdot \tau(a)$$

$$= \sum_{(a,k)\in A \times R} \pi_u(a,k)(k \cdot \tau(a)) = \tau_\pi(\pi_u) = \tau_\pi(\delta_\pi^*(u, u)) = \left(\left[\mathcal{A}_\phi\right], u\right),$$

and hence, $\mathcal{A}_\phi$ is $r$-equivalent to $\mathcal{A}$. □

Algorithm 9.2 (Construction of the cdwfa $\mathcal{A}_\phi$). The input of this algorithm is a weighted finite automaton $\mathcal{A} = (A, \delta, \sigma, \tau)$ over $X$ and $K$, which satisfies the properties of Theorem 9.1. The output is the accessible crisp-deterministic weighted finite automaton $\mathcal{A}_\phi = (A_\delta, \delta_\phi, \phi_\tau, \tau_\phi)$ which is $r$-equivalent to $\mathcal{A}$.

The procedure is to construct the transition tree of $\mathcal{A}_\phi$ directly from $\mathcal{A}$. This is done in the same way as in Algorithm 6.4 except that we replace: in (A1): $\sigma$ by $\phi_\tau$ in (A3) and (A4): $\sigma$ and $\sigma_\text{ux}$ by, resp., $\phi_u$ and $\phi_\text{ux}$; in (A4): $\tau_\phi$ by $\tau_\phi$ and (13) by (37); in (A5): $\mathcal{A}_\phi$ by $\mathcal{A}_\phi$.

Since $\mathcal{A}$ has the mentioned property that the set $R = (\text{Im}(\sigma) \cdot T) \setminus \{0\}$ is finite and every element of $R \cdot \text{Im}(\tau)$ is additively idempotent, where $T$ denotes the submonoid of $(K, \cdot, 1)$ generated by the set $\text{Im}(\delta)$, the algorithm terminates in a finite number of steps. In particular, if $K$ is additively idempotent and multiplicatively locally finite, then every wfa over $X$ and $K$ has this property. On the other hand, if $K$ is additively idempotent but not multiplicatively locally finite, then the algorithm does not necessarily terminate in a finite number of steps, but it can happen in some cases.

Algorithm 9.3 (Computing $\phi_\text{ux}$ from $\phi_u$). This procedure computes $\phi_\text{ux}$ directly from $\phi_u$, for every $u \in X'$ and $x \in X$:

1. Initialization: $\phi_\text{ux} = \emptyset$.
2. For all $(a,k) \in \phi_u$ and $b \in A$ such that $k \cdot \delta_x(a, b) \neq 0$ we extend $\phi_\text{ux}$ by

$$\phi_\text{ux} = \phi_\text{ux} \cup \{(b, k \cdot \delta_x(a, b))\}.$$

According to (35), at the end of this procedure we obtain the correct set $\phi_\text{ux}$.

10. Relationships between finiteness of $\mathcal{A}_N$ and $\mathcal{A}_M$, and definability of $\mathcal{A}_\pi$

In this section we examine relationships between finiteness of $\mathcal{A}_N$ and $\mathcal{A}_M$, and definability of $\mathcal{A}_\pi$. First we prove two results which establish certain relationships between the cardinalities of $\mathcal{A}_N$ and $\mathcal{A}_M$, and the cardinalities of $\mathcal{A}_N$ and $\mathcal{A}_\pi$.

Let $\mathcal{A} = (A, \delta, \sigma, \tau)$ be a wfa over $X$ and $K$. The weighted transition function $\delta$ is called deterministic if for every $x \in X$ and every $a \in A$ there exists at most one $a' \in A$ such that $\delta_x(a, a') \neq 0$ (cf. [11]).

Proposition 10.1. Let $\mathcal{A} = (A, \delta, \sigma, \tau)$ be a weighted finite automaton over $X$ and $K$.

(a) If $\sigma$ is crisp-deterministic, or $K$ is a semiring, then $|\mathcal{A}_N| \leq |\mathcal{A}_M|$. 

20
If $\delta$ is deterministic and the submonoid $T$ of $(K^-, 1)$ generated by the set $\text{Im}(\delta)$ is finite, then $\mathcal{A}_M$ is finite.

(c) If $\delta$ is crisp-deterministic, then both $\mathcal{A}_N$ and $\mathcal{A}_M$ are finite and $|\mathcal{A}_N| \leq |\mathcal{A}_M|$. Proof. (a) If $\sigma$ is crisp-deterministic, then by Theorem 5.1 we obtain that $\sigma_u = \sigma \cdot \delta_u$, for every $u \in X^*$. The same family of equations is fulfilled if $K$ is a semiring, since in this case the matrix and matrix-vector products (whenever defined) are associative. Define a function $\phi : A_M \rightarrow A_N$ by $\phi(\delta_u) = \sigma_u$, for every $u \in X^*$. If $\delta_u = \delta_v$, for some $u, v \in X^*$, then $\phi(\delta_u) = \sigma_u = \sigma \cdot \delta_u = \sigma \cdot \delta_v = \phi(\delta_v)$, and therefore, $\phi$ is a well-defined function. It is clear that $\phi$ is surjective, so we obtain that $|\mathcal{A}_N| \leq |\mathcal{A}_M|$. (b) If $\delta$ is deterministic, then for each $u \in X^*$ we have that $\delta_u : A \times A \rightarrow T$. Since $T$ is finite, then $A_M$ is also finite. Hence, $\mathcal{A}_M$ is finite.

(c) If $\delta$ is crisp-deterministic, then $T \subseteq \{0, 1\}$, and according to (b), $\mathcal{A}_M$ is finite. By Theorem 5.1 it follows that $\sigma_u = \sigma \cdot \delta_u$, for every $u \in X^*$, and as in the proof of (a) we obtain that $\mathcal{A}_N$ is also finite and $|\mathcal{A}_N| \leq |\mathcal{A}_M|$. □

Notice that the assertion (c) of the preceding theorem generalizes and simplifies the proof of (i)$\Rightarrow$(ii) of Theorem 3.2 of [25].

Theorem 10.2. Let $\mathcal{A} = (A, \delta, \sigma, \tau)$ be a weighted finite automaton over an alphabet $X$ and a right distributive strong bimonoid $K$, having the property that the submonoid $T$ of $(K^-, 1)$ generated by the set $\text{Im}(\delta)$ is finite.

(a) Let every element of $\text{Im}(\sigma)$ have finite additive order. Then $\mathcal{A}_n$ and $\mathcal{A}_m$ are finite and $|\mathcal{A}_N| \leq |\mathcal{A}_M|$. (b) Let every element of $\text{Im}(\delta)$ have finite additive order. Then $\mathcal{A}_M$ is finite.

Proof. (a) By the finiteness of $T$ it follows that $R = (\text{Im}(\sigma) \cdot T) \setminus \{0\}$ is also finite. If $s \in \text{Im}(\sigma)$ and $k \in K$, then $sk$ has finite additive order by the assumption of (a) and right distributivity of $K$, and consequently, any element of $R \cdot \text{Im}(\tau)$ has a finite additive order. Hence, the assumptions of Theorem 8.2 are satisfied, so $\mathcal{A}_n$ is defined and finite. Define a function $\psi : A_n \rightarrow A_N$ by $\psi(\pi_n) = \sigma_{u_n}$, for each $u \in X^*$.

If $u, v \in X^*$ such that $\pi_n = \pi_m$, then for every $a \in A$ we have that

$$\sigma_u(a) = \sum_{k \in K} p_u(a, k) k = \sum_{k \in K} \pi_u(a, k) k = \sum_{k \in K} \pi_v(a, k) k = \sum_{k \in K} p_v(a, k) k = \sigma_v(a),$$

by the right distributivity of $K$ and finiteness of $R$, so $\sigma_u = \sigma_v$. Therefore, $\psi$ is a well-defined function. It is clear that $\psi$ is surjective, and hence, $|\mathcal{A}_N| \leq |\mathcal{A}_m|$. (b) Let $S$ be a subsemigroup of $(K^-, 1)$ generated by the set $\text{Im}(\delta)$. By the assumption of the theorem, $S$ is finite, and by the same argument as in (a), every element of $S$ has finite additive order.

Set $i = \max \{i(k) | k \in S\} = \max \{i(k) | k \in \text{Im}(\delta)\}$ and $p = \text{lcm}\{p(k) | k \in S\} = \text{lcm}\{p(k) | k \in \text{Im}(\delta)\}$, where $i(k)$ is the index, and $p(k)$ is the period of $k$ in $(K^+, \cdot)$. For any $u \in X^*$ let us define functions $p_u' : A \times A \times T \rightarrow N$ and $\pi_u' : A \times A \times T \rightarrow \{0, 1, \ldots, i + p - 1\}$ as follows. For any $u \in X^*$, $a, b \in A$, and $k \in T$, let $p_u'(a, b, k)$ be the number of all paths from $a$ to $b$, labelled by $u$, with weight $k$, and let $p_u'(a, b, k) = 1$, if $a = b$ and $k = 1$, and $p_u'(a, b, k) = 0$, otherwise. Moreover, for each $u \in X^*$ let

$$\pi_u'(a, b, k) = \begin{cases} p_u'(a, b, k) & \text{if } p_u'(a, b, k) < i, \\ i + \left(\frac{p_u'(a, b, k) - i}{p}\right) & \text{if } p_u'(a, b, k) \geq i, \end{cases}$$

where $(p_u'(a, b, k) - i) \mod p$ is the remainder when $p_u'(a, b, k) - i$ is divided by $p$. Clearly, $\pi_u'(a, b, k) \equiv_p p_u'(a, b, k)$.

Set $A_{m'} = \{\pi_u' | u \in X^*\}$, and define a function $\psi' : A_n \rightarrow A_M$ by $\psi'(\pi_u') = \delta_{u'}$, for each $u \in X^*$. Let $u, v \in X^*$, $u \neq v$, such that $\pi_u = \pi_v$. If $u, v \in X^*$, then for all $a, b \in A$ we have that

$$\delta_{u}(a, b) = \sum_{k \in S} p_u'(a, b, k) k = \sum_{k \in S} \pi_u'(a, b, k) k = \sum_{k \in S} \pi_v'(a, b, k) k = \sum_{k \in S} p_v'(a, b, k) k = \delta_{v}(a, b),$$

and consequently, $\delta_{u} = \delta_{v}$. 21
Next, let \( u = \varepsilon \) or \( v = \varepsilon \). Without loss of generality we can assume that \( u \in X^+ \) and \( v = \varepsilon \). Consider arbitrary \( a, b \in A \) and \( k \in T \). If \( a = b \) and \( k = 1 \), then

\[
1 = p'_\varepsilon(a, b, k) = \pi'_\varepsilon(a, b, k) = \pi'_\varepsilon(a, b, k) \leq p'_\varepsilon(a, b, k),
\]

what implies that there is a non-empty path from \( a \) to \( b \) with weight \( k = 1 \), and hence \( 1 \in S \), and otherwise, if \( a \neq b \) or \( k \neq 1 \), then

\[
0 = p'_\varepsilon(a, b, k) = \pi'_\varepsilon(a, b, k) = \pi'_\varepsilon(a, b, k) = p'_\varepsilon(a, b, k).
\]

Now, if \( a = b \), then

\[
\delta_u(a, b) = \sum_{k \in S} p'_\varepsilon(a, b, k) k = p'_\varepsilon(a, b, 1)1 = \pi'_\varepsilon(a, b, 1)1 = 1 = \delta_\varepsilon(a, b),
\]

and if \( a \neq b \), then

\[
\delta_u(a, b) = \sum_{k \in S} p'_\varepsilon(a, b, k) k = 0 = \delta_\varepsilon(a, b),
\]

and hence, \( \delta_u = \delta_\varepsilon \), i.e., \( \delta_u = \delta_\varepsilon \).

Therefore, we conclude that \( \psi' \) is a well-defined surjective function, whence \( |A_M| \leq |A'_M| \) and \( A_{\pi'} \) is finite with \( |A_{\pi'}| = (i + p)|A^2|\). \( \Box \)

It can be similarly proved that \( A_{\pi} \) and \( A_{\pi'} \) are finite and \( |A_{\pi'}| \leq |A_{\pi}| \) whenever \( K \) is right distributive and \( A_{\pi} \) is defined (i.e., \( R \) is finite and every element of \( \text{Im}(\sigma) \) has finite additive order). Note also that if \( K \) is right distributive and 1 has finite additive order, then each element of \( K \) has finite additive order.

Let \( A = (A, \delta, \sigma, \tau) \) be a weighted finite automaton over an alphabet \( X \) and a strong bimonoid \( K \). If \( K \) is weakly locally finite, then all three automata \( A_{\pi}, A_M, \) and \( A_{\pi} \) are finite. Moreover, the following is true.

**Proposition 10.3.** Let \( K \) be a strong bimonoid.

(a) If for every wfa \( A \) over \( K \), \( A_{\pi} \) and \( A_{M} \) are finite, then \( K \) is bi-locally finite.

(b) Let \( K \) be right distributive. Then in (a) we have equivalence.

**Proof.** (a) The automata \( M \) and \( M' \) constructed in the proof of Lemma 12 of [11] have crisp-deterministic initial states. Hence the \( \text{i}\)-behavior of \( M \) coincides with the \( \text{i}\)-behavior of \( M' \), and correspondingly for \( M' \), by our Theorem 5.1. If their Nerode or Myhill automata are finite, then the \( \text{i}\)-behaviors of \( M \) and \( M' \) have finite image. By the argument in the proof of Lemma 12 of [11], the claim about \( K \) follows.

(b) Let \( K \) be bi-locally finite, and \( A \) a wfa over \( K \). Since \( K \) is bi-locally finite and right distributive, \( K \) is weakly locally finite. Hence the states of \( A_{\pi} \) and \( A_{M} \) are contained respectively in finite sets \( W^A \) and \( W^{A \times A} \), where \( W \) is the weak closure of the set \( \text{Im}(\sigma) \cup \text{Im}(\delta) \). \( \Box \)

Next we give examples which demonstrate that finiteness of \( A_{\pi} \) and \( A_{M} \), and definability of \( A_{\pi} \), are completely independent of each other, i.e., for a wfa \( A \) all combinations of finiteness/infiniteness of \( A_{\pi} \) and \( A_{M} \), and definability/undefinability of \( A_{\pi} \) are possible.

For a wfa \( A = (A, \delta, \sigma, \tau) \) over \( X \) and \( K \), the finiteness of \( A_{\pi} \) and \( A_{M} \) depends only on \( \delta \) and \( \sigma \), and if we assume that \( 1 \in \text{Im}(\tau) \subseteq [0, 1] \), then the assumptions that \( R = (\text{Im}(\sigma) \cdot T) \setminus \{0\} \) is finite and every element of \( R \cdot \text{Im}(\tau) \) has finite additive order also depend only on \( \delta \) and \( \sigma \). For that reason, in the sequel \( A = (A, \delta, \sigma, \tau) \) will be a wfa over an alphabet \( X = \{x\} \) and a strong bimonoid \( K \), with \( A \) having two states and \( \tau \) being any final weight vector satisfying the above condition \( 1 \in \text{Im}(\tau) \subseteq [0, 1] \), and \( K, \sigma \) and \( \delta \) will be specified separately.
Example 10.4. Let $K = (\mathbb{N}, +, \cdot, 0, 1)$, the semiring of natural numbers, and $\sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

(a) Put $\delta_x = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Then $\sigma \cdot \delta_x = \sigma$, so $\mathcal{A}_N$ is finite. Clearly $\mathcal{A}_M$ is finite.

(b) Put $\delta_x = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$. Then $\sigma \cdot \delta_x = \sigma$, so $\mathcal{A}_N$ is finite, and $\delta_{\pi^n} = \begin{bmatrix} 1 & 0 \\ n & 1 \end{bmatrix}$ for every $n \in \mathbb{N}$, so $\mathcal{A}_M$ is infinite.

(c) Put $\delta_x = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$. Then $\sigma_{\pi^n} = \begin{bmatrix} 2^n & 0 \\ 0 & 0 \end{bmatrix}$, for every $n \in \mathbb{N}\setminus\{0\}$, so $\mathcal{A}_N$ and $\mathcal{A}_M$ are infinite.

In each case, the set $K = (\text{Im}(\sigma) \cdot T) \setminus \{0\}$ is finite, but any of its elements has infinite additive order, and we conclude that the assumptions of Section 8 for $\mathcal{A}_n$ are not satisfied, and $\mathcal{A}_n$ is not defined.

Example 10.5. Let $K = (\mathbb{N}, \oplus, \odot, 0, 1)$ be the strong bimonoid with $a \oplus 1 = 1$ for each $a \in \mathbb{N}$, in particular $1 \oplus 1 = 1, a \oplus b = a + b$ (usual sum) if $a, b \geq 2$, and $\cdot$ is the usual multiplication in $\mathbb{N}$. Put

$$\sigma = \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}, \quad \delta_x = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$ 

Then $\delta_x \cdot \delta_x = \delta_x$ as $1 \oplus 1 = 1$, so $\mathcal{A}_M$ is finite, and $\sigma_{\pi^n} = \begin{bmatrix} 2^{n+1} & 2^{n+1} \\ 0 & 0 \end{bmatrix}$, for each $n \in \mathbb{N}$, so $\mathcal{A}_N$ is infinite.

As in Example 10.4 we obtain that the assumptions of Section 8 are not fulfilled, i.e., $\mathcal{A}_n$ is not defined.

Example 10.6. (cf. Example 25 of [11]) Let $K = (\mathbb{N}, \oplus, \odot, 0, 1)$ be the strong bimonoid with operations $\oplus$ and $\odot$ defined as follows. First, let $0 \oplus a = a, 0 \odot a = 0$, and $1 \odot a = a$ for every $a \in \mathbb{N}$. If $a, b \in \mathbb{N}\setminus\{0\}$ with $a \leq b$, we put (with $+$ being the usual addition on $\mathbb{N}$)

$$a \oplus b = \begin{cases} b, & \text{if } b \text{ is even;} \\ b + 1, & \text{if } b \text{ is odd.} \end{cases}$$

If $a, b \in \mathbb{N}\setminus\{0, 1\}$ with $a \leq b$, let

$$a \odot b = \begin{cases} b + 1, & \text{if } b \text{ is even;} \\ b, & \text{if } b \text{ is odd.} \end{cases}$$

(a) Put $\sigma = \begin{bmatrix} 2 & 2 \\ 1 & 0 \end{bmatrix}, \quad \delta_x = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$. Then for every $n \in \mathbb{N}\setminus\{0\}$ we have that

$$\sigma_{\pi^n} = \begin{bmatrix} 2n + 2 & 2n + 2 \\ 2n & 2n \end{bmatrix}, \quad \delta_{\pi^n} = \begin{bmatrix} 2n & 2n \\ 2n & 2n \end{bmatrix},$$

and hence, $\mathcal{A}_N$ and $\mathcal{A}_M$ are infinite.

(b) Put $\sigma = \begin{bmatrix} 1 & 0 \\ 2 & 2 \end{bmatrix}, \quad \delta_x = \begin{bmatrix} 2 & 0 \\ 2 & 2 \end{bmatrix}$. Then $\sigma_x = \begin{bmatrix} 2 & 0 \\ 2 & 2 \end{bmatrix}, \sigma_{\pi^n} = \begin{bmatrix} 3 & 0 \\ 2n & 3 \end{bmatrix}$, so $\mathcal{A}_N$ is finite. But, $\delta_{\pi^n} = \begin{bmatrix} 3 & 0 \\ 2n & 3 \end{bmatrix}$, for every $n \in \mathbb{N}, n \geq 2$, so $\mathcal{A}_M$ is infinite.

Since $K$ is bi-locally finite, in both cases we have that the assumptions of Section 8 for $\mathcal{A}_n$ are satisfied, and therefore, $\mathcal{A}_n$ is defined (and it is finite).

Example 10.7. We combine the idea of Example 10.5 with Example 10.6 (a).

Let $K = (\mathbb{N}, \oplus, \odot, 0, 1)$ where $a \oplus 1 = 1$ for each $a \in \mathbb{N}$, $a \odot b$ is defined as in Example 10.6, if $a, b \geq 2$; $2 \odot 2 = 2, a \odot b$ is defined as in Example 10.6 if $a \geq 3$ or $b \geq 3$. Then $K$ is a strong bimonoid and bi-locally finite, so $\mathcal{A}_n$ is defined. Put

$$\sigma = \begin{bmatrix} 4 & 4 \\ 2 & 2 \end{bmatrix}, \quad \delta_x = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}.$$
Then $\delta_x \cdot \delta_x = \delta_x$ by $2 \odot 2 = 2 \oplus 2 = 2$, so $\mathcal{A}_M$ is finite. On the other hand, $\sigma_{xy} = [2n + 4 \quad 2n + 4]$, for every $n \in \mathbb{N}$, so $\mathcal{A}_N$ is infinite.

The remaining case, when $\mathcal{A}_N$ and $\mathcal{A}_M$ are finite, and $\mathcal{A}_r$ is defined (and hence finite), will be considered in examples given in the next section.

11. Computational examples

**Example 11.1.** Consider the strong bimonoid $K = ([0,1], V_L, \Delta_L, 0, 1)$, where $V_L$ is the Łukasiewicz t-conorm and $\Delta_L$ is the Łukasiewicz t-norm, i.e., $x V_L y = \min(x + y, 1)$ and $x \Delta_L y = \max(x + y - 1, 0)$, for all $x, y \in [0,1]$. It is well-known that this bimonoid is not distributive. Let $\mathcal{A} = (A, \delta, \sigma, \tau)$ be a weighted finite automaton over an alphabet $X = \{x, y\}$ and $K$ given by the graph in Figure 2.

In other words, we have that

$$\sigma = \begin{bmatrix} 1 & 1 \\ 0 & \frac{1}{2} \end{bmatrix}, \quad \delta_x = \begin{bmatrix} 0 & \frac{1}{2} \\ 0 & 1 \end{bmatrix}, \quad \delta_y = \begin{bmatrix} 0 & 1 \\ 0 & \frac{1}{2} \end{bmatrix}, \quad \tau = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{38}$$

The transition trees and the graphs of the Nerode automaton $\mathcal{A}_N$, the Myhill automaton $\mathcal{A}_M$, and the cdwfa $\mathcal{A}_r$ are shown in Figures 3, 4, and 5, respectively, where $\blacksquare$ are closure marks.

These cdwfa are not only different, but they also have different behaviors. Namely, we have that

$$(\llbracket \mathcal{A}_N \rrbracket, xy) = (\llbracket \mathcal{A} \rrbracket, xy)\neq \frac{1}{2}, \quad (\llbracket \mathcal{A}_M \rrbracket, xy) = (\llbracket \mathcal{A} \rrbracket, xy) = (\llbracket \mathcal{A}_r \rrbracket, xy) = 0.$$

Therefore, the initial algebra semantics differs both from transition and run semantics.

It is worth noting that an equivalence relation $\mu$ on $A_M$ with classes $\{\delta_1\}, \{\delta_x\}, \{\delta_y\}, \{\delta_{xy}\}, \{\delta_{x^2}\}$ is a congruence on $\mathcal{A}_M$ and the related factor automaton is isomorphic to $\mathcal{A}_N$. According to Theorem 5.1 of [19] we have that $\llbracket \mathcal{A}_M \rrbracket = \llbracket \mathcal{A}_N \rrbracket$, what yields $\llbracket \mathcal{A} \rrbracket_r = \llbracket \mathcal{A}_r \rrbracket_r$, i.e., the transition and run semantics of $\mathcal{A}$ coincide. Moreover, by Theorems 5.2 and 5.3 of [19], $\mathcal{A}_N$ is the minimal cdwfa of the series $\llbracket \mathcal{A} \rrbracket_r$, and also, $\mathcal{A}_N$ is the minimal cdwfa of the series $\llbracket \mathcal{A} \rrbracket_r$.

**Example 11.2.** Let us change the final weight vector $\tau$ from (38) to

$$\tau = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \tag{39}$$

i.e., let a wfa $\mathcal{A}$ be given as in Figure 6. Then the automata $\mathcal{A}_N$, $\mathcal{A}_M$, and $\mathcal{A}_r$ are those shown in Figure 7. In this case we have that

$$(\llbracket \mathcal{A}_N \rrbracket, x) = (\llbracket \mathcal{A} \rrbracket, x) = (\llbracket \mathcal{A}_M \rrbracket, xy) = (\llbracket \mathcal{A} \rrbracket, xy) = \frac{1}{2}, \quad (\llbracket \mathcal{A}_r \rrbracket, x) = (\llbracket \mathcal{A} \rrbracket_r, x) = 0.$$

Therefore, here we have an example of a wfa in which run semantics differs both from initial algebra semantics and transition semantics.

Let us observe that none of the automata $\mathcal{A}_N$, $\mathcal{A}_M$, and $\mathcal{A}_r$ is minimal. Namely, merging states $\sigma_{xy}$ and $\sigma_{x^2}$ in the automaton $\mathcal{A}_N$, and merging states $\delta_x$ and $\delta_y$, and states $\delta_{xy}$, $\delta_{x^2}$, and $\delta_{x^2}$ in the automaton $\mathcal{A}_M$, we obtain a minimal automaton shown in Figure 8 (left), and merging states $\pi_x$, $\pi_{xy}$, and $\pi_{x^2}$ in the automaton $\mathcal{A}_r$, we obtain a minimal automaton shown in Figure 8 (right).

![Figure 2: A wfa $\mathcal{A}$ from Example 11.1.](image-url)
Figure 3: The transition tree and the graph of the Nerode automaton $\mathcal{A}_N$ of the wfa $\mathcal{A}$ given in Figure 2.

Figure 4: The transition tree and the graph of the Myhill automaton $\mathcal{A}_M$ of the wfa $\mathcal{A}$ given in Figure 2.

Figure 5: The transition tree and the graph of the cdwfa $\mathcal{A}_c$ constructed starting from the wfa $\mathcal{A}$ given in Figure 2.

Figure 6: A wfa $\mathcal{A}$ from Example 11.2.
Example 11.3. Let us change the final weight vector $\tau$ from (38) to

$$\tau = \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \tag{40}$$

i.e., let a wfa $\mathcal{A}$ be given as in Figure 9. Then the cdwfa $\mathcal{A}_N$, $\mathcal{A}_M$, and $\mathcal{A}_\pi$ are those shown in Figure 10.

Evidently, none of these automata is minimal, and minimizing any of them we obtain the same automaton $\mathcal{A}_{\min}$, which is shown in Figure 11. By this we conclude that $\llbracket \mathcal{A}_N \rrbracket = \llbracket \mathcal{A}_M \rrbracket = \llbracket \mathcal{A}_\pi \rrbracket = \llbracket \mathcal{A}_{\min} \rrbracket$, and hence, $\llbracket \mathcal{A} \rrbracket = \llbracket \mathcal{A}_N \rrbracket = \llbracket \mathcal{A}_M \rrbracket = \llbracket \mathcal{A}_\pi \rrbracket$. Therefore, initial, transition and run semantics of $\mathcal{A}$ coincide, although $K$ is not (left, right) distributive, but the three determinization methods developed here result in three different automata, having different numbers of states.

Example 11.4. Consider the additively idempotent semiring $K = ([0,1], \max, \cdot, 0, 1)$, where $\cdot$ is the ordinary multiplication of real numbers. The semiring $K$ is not locally finite, since it is not multiplicatively locally finite. Let $\mathcal{A} = (A, \delta, \sigma, \tau)$ be a wfa over the alphabet $X = \{x\}$ and $K$ given by the graph in Figure 12 (left).
The Nerode automaton \( \mathcal{A}_N \) has two states, and it is shown in Figure 12 (right). On the other hand, the Myhill automaton \( \mathcal{A}_M \) is infinite (see Example 11.2 [19]), since for every \( n \in \mathbb{N} \setminus \{0\} \) we have that

\[
\delta_{x \cdot 2n-1} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2^{2n-1}} \end{bmatrix}, \quad \delta_{x \cdot 2n} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2^{2n}} \end{bmatrix}.
\]

In the notation of Sections 8 and 9, the submonoid \( T \) of \((\mathcal{K} \cdot, 1)\) generated by the set \( \text{Im(}\delta)\) is infinite, and consequently, the sets \( R \) and \( A \times R \) are infinite, but in spite of that, we can construct the cdwa \( \mathcal{A}_\pi \) (i.e., \( \mathcal{A}_\nu \)), and it is finite. Namely, it can be easily verified that \( \mathcal{A}_\nu \) has two states and it is isomorphic to \( \mathcal{A}_N \).

It is worth noting that the same cdwa \( \mathcal{A}_N, \mathcal{A}_M, \) and \( \mathcal{A}_\pi \) are obtained if we replace the maximum t-conorm \( \lor \) by the Łukasiewicz t-conorm \( \lor_L \). In this case \( K \) is not additively idempotent, and need not be a semiring, since any t-norm distributes over the maximum, but not necessarily over other t-conorms.